



PT New ligands for prion proteins, useful for detection or removal or  
PT prions and for treating prion-associated diseases, recognize a specific  
PT octapeptide motif -

Claim 1; Page 33; 47pp; English.

CC The invention relates to a ligand of less than 6 kD that binds to a  
CC polypeptide containing the sequence GYTRGYSINPRQISGLYSGLY (A) or an  
CC analogue that is the retro-inverse isomer of (A). The sequence A is  
CC an octapeptide motif from the prion protein (PrP). The ligands are  
CC identified by binding assays with the peptide (A) or peptides containing  
CC (A). The ligands are used for detecting prion proteins (or prions) in  
CC biological or environmental samples, e.g. for diagnosis, also for  
CC removing them from samples, for treating or retarding development of  
CC prion-associated diseases, especially Creutzfeldt-Jacob diseases (in  
CC iatrogenic, new variant, familial or sporadic forms), but also  
CC Gerstmann-Strausler-Scheinker disease, fatal familial insomnia,  
CC scrapie, bovine or feline spongiform encephalopathy, transmissible mink  
CC or exotic ungulate encephalopathy, or chronic wasting disease. The  
CC present sequence is the PrP octapeptide A.

Sequence 8 AA:

Query Match 100.0%; Score 55; DB 23; Length 8;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8  
Db 1 GWGQPHGG 8

RESULT 2

ABB81629  
ID ABB81629 standard; peptide: 11 AA.

XX ABB81629;

DT 25-SEP-2002 (first entry)

DE Prion mimetic peptide SEQ ID NO:1.

XX Prion mimetic peptide; degradation: detection; TSE; infection;  
KW transmissible spongiform encephalopathy; prion protein; sterilisation;  
KW immunisation; Creutzfeldt-Jacob disease; kuru; fatal familial insomnia;  
KW Gerstmann-Strausler-Scheinker syndrome; chronic wasting disease;  
KW bovine spongiform encephalopathy; feline spongiform encephalopathy;  
KW scrapie; transmissible mink encephalopathy.

Synthetic.

PN WO200253723-A2.

PD 11-JUL-2002.

PF 08-JAN-2002; 2002WO-GB00052.

PR 08-JAN-2001; 2001GB-0000420.

PR 26-FEB-2001; 2001GB-0004696.

PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.

PI Raven NDH;

DR WPI; 2002-557743/59.

PT Inactivating transmissible spongiform encephalopathy (TSE) agent such  
PT as Creutzfeldt-Jacob disease, scrapie, kuru or  
PT Gerstmann-Strausler-Scheinker syndrome involves exposing agent to  
PT thermostable proteolytic enzyme -  
XX Example; Page 19; 41pp; English.

CC The present invention describes a method (M1) for inactivating a  
CC transmissible spongiform encephalopathy (TSE) agent comprising exposing  
CC the TSE agent to a thermostable proteolytic enzyme. Also described:  
CC (1) a composition (I) for inactivating a TSE agent, comprising a  
CC thermostable proteolytic enzyme; (2) an antibody (II) specific for a  
CC prion dimer which does not bind to a prion monomer; and (3) a purified  
CC prion dimer. (M1) is useful for inactivating a TSE agent such as a prion.

CC A TSE agent is Creutzfeldt-Jacob disease or its variant, kuru, fatal  
CC familial insomnia, Gerstmann-Strausler-Scheinker syndrome, bovine  
CC spongiform encephalopathy, scrapie, feline spongiform encephalopathy,  
CC chronic wasting disease or transmissible mink encephalopathy. (I) is  
CC useful for sterilising material contaminated with the TSE agent. A prion  
CC dimer is useful for examining a sample infected with or suspected to be  
CC infected by a prion protein, and for detecting prion infectivity, by  
CC detecting a prion dimer in the sample. A prion dimer is useful for  
CC producing (II), by immunising an animal with a prion dimer, obtaining its  
CC extract which contains (II), and isolating (II) from the extract. The  
CC method comprises obtaining an antibody preparation containing antibodies  
CC which bind a prion dimer, and removing (II) from the preparation. (M1)  
CC and (I) are useful for inactivating TSE agents in potentially  
CC contaminated clinical waste and culled animal material. (M1) is useful  
CC for sterilising larger surface areas of apparatus, operating tables or  
CC even walls of rooms. The present sequence represents a prion mimetic  
CC peptide which is used in an example from the present invention in the  
CC preparation of antibodies including dimer preferential antibodies.

Sequence 11 AA:

Query Match 100.0%; Score 55; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.027;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8  
Db 3 GWGQPHGG 10

RESULT 3

ABB84521  
ID ABB84521 standard; peptide: 14 AA.

XX ABB84521;

DT 05-SEP-2001 (first entry)

DE Octa-peptide motif repeated in human prion protein.

XX Prion protein; subacute transmissible spongiform encephalopathy; ESST;  
KW Creutzfeld-Jacob disease; bovine spongiform encephalopathy; scrapie.

XX Homo sapiens.

PN WO200135104-A1.

PD 17-MAY-2001.

PF 13-NOV-2000; 2000WO-FR03159.

PR 12-NOV-1999; 99FR-0014242.

PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.

PI Deslys J, Comoy E, Grassi J;

DR WPI; 2001-408079/43.

PT Diagnosis of prion diseases, by treatment with proteinase K and  
PT detecting retention of octapeptide repeat motifs, including  
PT differentiation between prion strains -

XX Example 1; Page 13; 51pp; French.

XX The present sequence represents an octa-peptide repeat motif of a

CC prion protein. The specification describes a method for the diagnosis  
CC of subacute transmissible spongiform encephalopathy (ESST) caused by a  
CC strain of unconventional transmissible agent. The method comprises  
CC detecting abnormal prion protein in a biological sample. The sample is  
CC treated with at least one proteinase K while retaining at least some of  
CC the octapeptide repeats in the prion protein, then treatment with a  
CC ligand for octapeptide repeats and detecting any formation of a complex.  
CC The method is used to diagnose ESST, particularly (new variant)  
CC Creutzfeldt-Jakob diseases, bovine spongiform encephalopathy and  
CC scrapie.

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 55; DB 22; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.034;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8  
| | | | | | | |  
2 GWGQPHGG 9

RESULT 4  
AAU11825

ID AAU11825 standard; peptide: 16 AA.

XX AC AAU11825;

DT 26-MAR-2002 (first entry)

XX DE Prion protein, PrP, octapeptide motif containing peptide.

KW Prion protein; PrP; ligand: octapeptide motif; scrapie;  
KM prion-associated disease; Creutzfeldt-Jakob disease;  
KM Gerstmann-Straussler-Scheinker disease; fatal familial insomnia;  
KM feline spongiform encephalopathy; bovine spongiform encephalopathy;  
KM transmissible mink encephalopathy; exotic ungulate encephalopathy;  
KM chronic wasting disease.

XX OS Mammalia.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "Gly is acetylated"

FT Modified-site 16 /note= "Gly is amidated"

XX FT

XX WO200177687-A2.

XX PD 18-OCT-2001.

XX PF 05-APR-2001; 2001WO-US11150.

XX PR 05-APR-2000; 2000US-0543188.

XX PA (VITE-) VI TECHNOLOGIES INC.

XX PI Hammond DJ, Wiltshire VR, Carbonell R, Shen H;

XX DR WPI: 2002-061944/08.

XX PT New ligands for prion proteins, useful for detection or removal or

XX PS octapeptide motif -

XX CC Claim 1; Page 33; 47pp; English.

XX The invention relates to a ligand of less than 6 kD that binds to a  
XX polypeptide containing the sequence GlyTyrGlyGlnProHisGlyGly (A) or an  
XX analogue that is the retro-inverse isomer of (A). The sequence A is  
XX an octapeptide motif from the prion protein (PrP). The ligands are  
XX identified by binding assays with the peptide (A) or peptides containing  
XX (A). The ligands are used for detecting prion proteins (or prions) in

CC biological or environmental samples, e.g. for diagnosis, also for  
CC removing them from samples, for treating or retarding development of  
CC prion-associated diseases, especially Creutzfeldt-Jakob diseases (in  
CC iatrogenic, new variant, familial or sporadic forms), but also  
CC Gerstmann-Straussler-Scheinker disease, fatal familial insomnia,  
CC scrapie, bovine or feline spongiform encephalopathy, transmissible mink  
CC or exotic ungulate encephalopathy, or chronic wasting disease. The  
CC present sequence is a peptide containing the PrP octapeptide A used  
CC to isolate the ligands of the invention.

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 55; DB 23; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.038;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8  
| | | | | | | |  
DB 1 GWGQPHGG 8

RESULT 5  
AAU11856

ID AAU11856 standard; peptide: 16 AA.

XX AC AAU11856;

DT 26-MAR-2002 (first entry)

XX DE Prion protein, PrP, octapeptide motif containing peptide variant #1.

KW Prion protein; PrP; ligand: octapeptide motif; scrapie;  
KM prion-associated disease; Creutzfeldt-Jakob disease;  
KM Gerstmann-Straussler-Scheinker disease; fatal familial insomnia;  
KM feline spongiform encephalopathy; bovine spongiform encephalopathy;  
KM transmissible mink encephalopathy; exotic ungulate encephalopathy;  
KM chronic wasting disease.

XX OS Mammalia.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "Gly is acetylated"

FT Modified-site 16 /note= "Gly is amidated"

XX FT

XX WO200177687-A2.

XX PD 18-OCT-2001.

XX PF 05-APR-2001; 2001WO-US11150.

XX PR 05-APR-2000; 2000US-0543188.

XX PA (VITE-) VI TECHNOLOGIES INC.

XX PI Hammond DJ, Wiltshire VR, Carbonell R, Shen H;

XX DR WPI: 2002-061944/08.

XX PT New ligands for prion proteins, useful for detection or removal or

XX PS octapeptide motif -

XX PS Disclosure: Page 10; 47pp; English.

XX The invention relates to a ligand of less than 6 kD that binds to a  
XX polypeptide containing the sequence GlyTyrGlyGlnProHisGlyGly (A) or an  
XX analogue that is the retro-inverse isomer of (A). The sequence A is  
XX an octapeptide motif from the prion protein (PrP). The ligands are  
XX identified by binding assays with the peptide (A) or peptides containing  
XX (A). The ligands are used for detecting prion proteins (or prions) in

CC biological or environmental samples, e.g. for diagnosis, also for  
CC removing them from samples, for treating or retarding development of  
CC prion-associated diseases, especially Creutzfeldt-Jacob diseases (in  
CC iatrogenic, new variant, familial or sporadic forms), but also  
CC Gerstmann-Strausler-Scheinker disease, fatal familial insomnia,  
CC scrapie, bovine or feline spongiform encephalopathy, transmissible mink  
CC or exotic ungulate encephalopathy, or chronic wasting disease. The  
CC present sequence is a peptide containing the PrP octapeptide A  
CC (containing the octapeptide in a different permutation to that shown in  
CC AA011825) used to isolate the ligands of the invention.

XX SQ Sequence 16 AA:

Query Match 100.0%; Score 55; DB 23; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.038; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWCQPHGC 8  
|||||||  
5 GWCQPHGC 12

RESULT 6

AA038026 AAR38026 standard; protein: 18 AA.

XX AC AAR38026;

XX DT 14-OCT-1993 (first entry)

XX DE Bovine prion protein region E #2.

XX KM Antigen; prion; protein; region; frame shift; repeat; mutation; PrPc;  
XX FSa; FSb; subfragment; antibody; treatment; spongiform encephalopathy;  
XX human; sheep; cattle; cellular binding; aggregation; mammal; scrapie;  
XX immune system; PrPsc; ratio-inverso peptide; enzymatic degradation;  
XX resistance.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 1 /note= "One or more residues or may be absent"

FT Misc-difference 2 /note= "May be absent"

FT Misc-difference 3 /note= "May be absent"

FT Misc-difference 4 /note= "May be absent"

FT Misc-difference 5 /note= "May be absent"

FT Misc-difference 14 /note= "May be absent"

FT Misc-difference 15 /note= "May be absent"

FT Misc-difference 16 /note= "May be absent"

FT Misc-difference 17 /note= "May be absent"

FT Misc-difference 18 /note= "May be absent"

XX PN W09311155-A.

XX PD 10-JUN-1993.

XX PF 03-DEC-1992; 92WO-GB02246.

XX PR 03-DEC-1991; 91GB-0025747.

XX PR 10-JUL-1992; 92GB-0014663.

XX PA (PROT-) PROTEUS MOLECULAR DESIGN LTD.

PI Fishleigh RV, Mee RP, Robson B;

XX DR WPI, 1993-196994/24.

XX PT New polypeptide(s) contg. antigenic site of prion protein -

XX PT useful for treatment and diagnosis of mammalian encephalopathies

XX PS e.g. Creutzfeldt-Jacob disease and kuru

XX PS Claim 23; Page 70; 82pp; English.

XX CC The sequences given in AAR38025-36 represent polypeptides derived from  
XX CC an antigenic site, region E, of a prion protein. Prion proteins  
XX CC comprise six regions of interest (A-F), and two related frame shift  
XX CC peptides sequences caused by a repeating section in region E having  
XX CC a nucleic acid coding sequence frame shift mutation of +1 (Fsa) or -1  
XX CC (Fsb) (see also AAR38037-38). These peptides and antibodies raised  
XX CC against these may be used to treat or prevent spongiform encephalopathy  
XX CC in humans, sheep or cattle. They can be used to block cellular binding  
XX CC and aggregation of prion proteins and to stimulate the mammalian immune  
XX CC system. These peptides may be used to distinguish between the normal  
XX CC form of prion protein (PrPc) and the scrapie-associated form (PrPsc).  
XX CC These peptides may include rare or synthetic amino acids or a ratio-  
XX CC inverso peptide modification to improve resistance to enzymatic  
XX CC degradation.

XX SQ Sequence 18 AA:

Query Match 100.0%; Score 55; DB 14; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.042; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWCQPHGC 8  
|||||||  
4 GWCQPHGC 11

RESULT 7

AA038026 AAR38026 standard; protein: 18 AA.

XX AC AAR38026;

XX DT 14-OCT-1993 (first entry)

XX DE Ovine prion protein region E #2.

XX KM Antigen; prion; protein; region; frame shift; repeat; mutation; PrPc;  
XX FSa; FSb; subfragment; antibody; treatment; spongiform encephalopathy;  
XX human; sheep; cattle; cellular binding; aggregation; mammal; scrapie;  
XX immune system; PrPsc; ratio-inverso peptide; enzymatic degradation;  
XX resistance.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 1 /note= "One or more residues or may be absent"

FT Misc-difference 2 /note= "May be absent"

FT Misc-difference 3 /note= "May be absent"

FT Misc-difference 4 /note= "May be absent"

FT Misc-difference 5 /note= "May be absent"

FT Misc-difference 14 /note= "May be absent"

FT Misc-difference 15 /note= "May be absent"

FT Misc-difference 16 /note= "May be absent"

FT Misc-difference 17 /note= "May be absent"



```

FT Misc-difference 18 /note= "One or more residue or may be absent"
FT
XX
PN WO9311155-A.
XX
PD 10-JUN-1993.
XX
PF 03-DEC-1992; 92WO-GB02246.
XX
PR 03-DEC-1991; 91GB-0025747.
PR 10-JUL-1992; 92GB-0014663.
XX
PA (PROT-) PROTEUS MOLECULAR DESIGN LTD.
XX
PI Fishleigh RV, Mee RP, Robson B;
XX
DR WPI: 1993-196994/24.
XX
PS New polypeptide(s) contg. antigenic site of prion protein -
PS e.g. Creutzfeld-Jacob disease and kuru
PS Claim 23; Page 70-71; 82pp; English.
XX
PS The sequences given in AAR38025-36 represent polypeptides derived from
XX an antigenic site, region E, of a prion protein. Prion proteins
XX comprise six regions of interest (A-F), and two related frame shift
XX peptides sequences caused by a repeating section in region E having
XX a nucleic acid coding sequence frame shift mutation of +1 (Fsa) or -1
XX (Fsb) (see also AAR38037-38). These peptides and antibodies raised
XX against these may be used to treat or prevent spongiform encephalopathy
XX in humans, sheep or cattle. They can be used to block cellular binding
XX and aggregation of prion proteins and to stimulate the mammalian immune
XX system. These peptides may be used to distinguish between the normal
XX form of prion protein (PrPc) and the scrapie-associated form (PrPsc).
XX These peptides may include rare or synthetic amino acids or a ratio-
XX inverse peptide modification to improve resistance to enzymatic
XX degradation.
XX
SQ Sequence 18 AA:
QY Query Match 100.0%; Score 55; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GWGQPHGG 8
|||||||
4 GWGQPHGG 11
RESULT 8
AAR38032
ID AAR38032 standard; protein; 18 AA.
XX
AC AAR38032;
XX
DT 14-OCT-1993 (first entry)
XX
DE Human prion protein region E #2.
XX
KW Antigen; prion; protein; region; frame shift; repeat; mutation; PrPc;
KW Fsa; Fsb; subfragment; antibody; treatment; spongiform encephalopathy;
KW human; sheep; cattle; cellular binding; aggregation; mammal; scrapie;
KW immune system; PrPsc; ratio-inverse peptide; enzymatic degradation;
KW resistance.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /note= "One or more residues or may be absent"
FT Misc-difference 2 /note= "May be absent"
FT

```

```

FT Misc-difference 3 /note= "May be absent"
FT
FT Misc-difference 4 /note= "May be absent"
FT
FT Misc-difference 5 /note= "May be absent"
FT
FT Misc-difference 14 /note= "May be absent"
FT
FT Misc-difference 15 /note= "May be absent"
FT
FT Misc-difference 16 /note= "May be absent"
FT
FT Misc-difference 17 /note= "May be absent"
FT
FT Misc-difference 18 /note= "May be absent"
FT
FT Misc-difference 18 /note= "One or more residue or may be absent"
FT
XX
PN WO9311155-A.
XX
PD 10-JUN-1993.
XX
PF 03-DEC-1992; 92WO-GB02246.
XX
PR 03-DEC-1991; 91GB-0025747.
PR 10-JUL-1992; 92GB-0014663.
XX
PA (PROT-) PROTEUS MOLECULAR DESIGN LTD.
XX
PI Fishleigh RV, Mee RP, Robson B;
XX
DR WPI: 1993-196994/24.
XX
PS New polypeptide(s) contg. antigenic site of prion protein -
PS useful for treatment and diagnosis of mammalian encephalopathies
PS e.g. Creutzfeld-Jacob disease and kuru
PS Claim 23; Page 71; 82pp; English.
XX
PS The sequences given in AAR38025-36 represent polypeptides derived from
XX an antigenic site, region E, of a prion protein. Prion proteins
XX comprise six regions of interest (A-F), and two related frame shift
XX peptides sequences caused by a repeating section in region E having
XX a nucleic acid coding sequence frame shift mutation of +1 (Fsa) or -1
XX (Fsb) (see also AAR38037-38). These peptides and antibodies raised
XX against these may be used to treat or prevent spongiform encephalopathy
XX in humans, sheep or cattle. They can be used to block cellular binding
XX and aggregation of prion proteins and to stimulate the mammalian immune
XX system. These peptides may be used to distinguish between the normal
XX form of prion protein (PrPc) and the scrapie-associated form (PrPsc).
XX These peptides may include rare or synthetic amino acids or a ratio-
XX inverse peptide modification to improve resistance to enzymatic
XX degradation.
XX
SQ Sequence 18 AA:
QY Query Match 100.0%; Score 55; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GWGQPHGG 8
|||||||
4 GWGQPHGG 11
RESULT 9
AAR38035
ID AAR38035 standard; protein; 26 AA.
XX
AC AAR38035;
XX
DT 14-OCT-1993 (first entry)
XX
DE Prion protein region E #2.

```

XX Antigen: prion; protein; region; frame shift; repeat; mutation; prpsc;  
 KW F5a; F5b; subfragment; antibody; treatment; spongiform encephalopathy;  
 KW human; sheep; cattle; cellular binding; aggregation; mammal; scrapie;  
 KW immune system; prpsc; ratio-inverso peptide; enzymatic degradation;  
 KW resistance.  
 OS Synthetic.  
 PN WO9311155-A.  
 XX 10-JUN-1993.  
 PD 03-DEC-1992; 92WO-GB02246.  
 PF 03-DEC-1991; 91GB-0025747.  
 PR 10-JUL-1992; 92GB-0014663.  
 XX (PROT-) PROTEUS MOLECULAR DESIGN LTD.  
 PA Fishleigh RV, Mee RP, Robson B;  
 DR WPI; 1993-196994/24.  
 PT New polypeptide(s) contg. antigenic site of prion protein -  
 PT useful for treatment and diagnosis of mammalian encephalopathies  
 PT e.g. Creutzfeldt-Jacob disease and kuru  
 XX Claim 24: Page 71: 82pp; English.  
 PS The sequences given in AAR38025-36 represent polypeptides derived from  
 CC an antigenic site, region E, of a prion protein. Prion proteins  
 CC comprise six regions of interest (A-F), and two related frame shift  
 CC peptides sequences caused by a repeating section in region E having  
 CC a nucleic acid coding sequence frame shift mutation of +1 (F5a) or -1  
 CC (F5b) (see also AAR38037-38). These peptides and antibodies raised  
 CC against these may be used to treat or prevent spongiform encephalopathy  
 CC in humans, sheep or cattle. They can be used to block cellular binding  
 CC and aggregation of prion proteins and to stimulate the mammalian immune  
 CC system. These peptides may be used to distinguish between the normal  
 CC form of prion protein (PrP<sup>C</sup>) and the scrapie-associated form (PrP<sup>Sc</sup>).  
 CC These peptides may include rare or synthetic amino acids or a ratio-  
 CC inverso peptide modification to improve resistance to enzymatic  
 CC degradation.  
 CC Sequence 26 AA:  
 SQ  
 Query Match 100.0%; Score 55; DB 14; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 0.06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GWCQPHGG 8  
 DB 7 GWCQPHGG 14  
 DE  
 RESULT 10  
 AAY07999  
 ID AAY07999 standard; protein: 42 AA.  
 AC AAY07999;  
 XX 08-JUL-1999 (first entry)  
 DT  
 XX Bovine prion protein derived peptide III.  
 DE  
 XX Prion protein; PrP; human; polyclonal antiserum; immunoassay;  
 KW detection; bovine; murine.  
 XX Bos taurus.  
 OS  
 XX DE19745443-A1.  
 PN  
 XX

PD 22-APR-1999.  
 XX  
 XX 15-OCT-1997; 97DE-1045443.  
 XX 15-OCT-1997; 97DE-1045443.  
 PR (HERZ/) HERZOG-MESMER A.  
 XX  
 XX Kiselev OI, Mesmer AH, Scheller A;  
 PI  
 DR WPI; 1999-255775/22.  
 XX  
 PT Diagnostic polyclonal antiserum specific for prion protein -  
 PT obtained by immunisation with metal-containing polypeptide  
 XX  
 PS Disclosure: Page 4; 12pp; German.  
 XX  
 CC This invention describes a novel process for producing a polyclonal  
 CC antiserum against a human or animal prion protein (PrP) which can be  
 CC used in immunoassays for detecting PrP's. The method comprises (a)  
 CC selecting a polypeptide that has a length of at least 10 amino acids and  
 CC has an amino acid sequence at least 70% homologous to that of human,  
 CC bovine or murine PrP in a region of at least 10 consecutive amino acids  
 CC (b) binding a metal to the polypeptide by reaction with a metal compound  
 CC and (c) injecting the metal-containing polypeptide into a host animal,  
 CC optionally together with adjuvants, to induce production of a polyclonal  
 CC antiserum.  
 CC  
 SQ Sequence 42 AA:  
 Query Match 100.0%; Score 55; DB 20; Length 42;  
 Best Local Similarity 100.0%; Pred. No. 0.094;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GWCQPHGG 8  
 DB 4 GWCQPHGG 11  
 DE  
 RESULT 11  
 AAW70280  
 ID AAW70280 standard; peptide: 178 AA.  
 AC AAW70280;  
 XX  
 DT 06-NOV-1998 (first entry)  
 XX  
 XX Peptide sequences used to raise antibodies against prion protein.  
 DE  
 XX Spinal cord; cattle; sheep; pig; bovine spongiform encephalopathy;  
 KW BSE; scrapie; transmissible spongiform encephalopathy; TSE;  
 KW immunological assay; scrapie prion protein; prpsc.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH 28. 61  
 FT Region /note= "Synthetic sequence used to raise antibody  
 FT against prpsc"  
 FT Region 90. 128  
 FT /note= "Synthetic sequence used to raise antibody  
 FT against prpsc"  
 XX  
 PN WO9835236-A2.  
 XX  
 PD 13-AUG-1998.  
 XX  
 PF 06-FEB-1998; 98WO-IE00007.  
 XX  
 PR 01-MAY-1997; 97IE-0000325.  
 PR 06-FEB-1997; 97IE-0000081.  
 PR 24-MAR-1997; 97IE-0000228.  
 XX

PA (ENFE-) ENFER TECHNOLOGY LTD.  
XX  
PI O'Connor M;  
XX  
DR WPI; 1998-447377/38.  
XX  
PT Detecting pathogenic prion(s) in animal carcasses - by reaction with  
PT specific labelled antibody; used to detect those carrying agents for  
PT bovine spongiform encephalopathy and scrapie  
XX  
PS Claim 3; Page 23; 25pp; English.  
XX  
CC The invention claims to provide a method for detecting transmissible  
CC spongiform encephalopathies (TSE) in animals and in animal carcasses.  
CC The method comprises of an immunological assay whereby the animal test  
CC sample is reacted with a labelled antibody against scrapie prion protein  
CC (prpsc) and the amount of bound labelled antibody is then detected. The  
CC anti-prpsc antibodies used in the assay are raised against fragments of  
CC the present synthetic peptide shown. The peptide fragments preferred  
CC by the inventors are shown in the features table. The method is claimed  
CC to be useful when applied to samples, particularly a cross-section of  
CC the spinal cord, from cattle, sheep and pig carcasses for detection of  
CC bovine spongiform encephalopathy (BSE) or scrapie.  
XX  
SQ Sequence 178 AA:  
  
Query Match 100.0%; Score 55; DB 19; Length 178;  
Best Local Similarity 100.0%; Pred. No. 0.37; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GWGDPHGG 8  
ID 77 GWGDPHGG 84  
DB  
  
RESULT 12  
AAB07316  
ID AAB07316 standard; protein; 208 AA.  
XX  
AC AAB07316;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Mouse prion protein sequence.  
XX  
DE Mouse prion protein; transmissible spongiform encephalopathy;  
KW bovine spongiform encephalopathy; BSE diagnosis; TSE; PrP.  
XX  
KW Mus sp.  
XX  
FH Key Location/Qualifiers  
FH Region 37..68  
FT /note="Repeat region consisting of tandem repeats  
FT of repeat unit: PHGGGWGQ (AAB07319)"  
FT Disulfide-bond 156..191  
FT Modified-site 208  
FT /note="C-terminal phospho-inositol glycolipid  
FT membrane anchor (-GPI)"  
XX  
XX WO200029850-A1.  
XX  
XX 25-MAY-2000.  
XX  
XX 27-OCT-1999; 99WO-FI00897.  
XX  
XX 17-NOV-1998; 98FI-0002481.  
XX  
XX (WALL-) WALLAC OY.  
XX (BBSR-) BBSRC OFFICE.  
XX  
XX Hope J, Barnard GJR, Birkett CR,  
XX  
XX WPI; 2000-387880/33.

XX  
PT Novel immunoassay for prion protein, used for the determination of  
PT transmissible spongiform encephalopathies in bovines -  
XX  
PS Disclosure; Page 41-42; 50pp; English.  
XX  
CC The present sequence is the mouse prion protein (PrP) sequence.  
CC Conversion of the normal cellular form of PrP into an aggregated,  
CC insoluble isoform is implicated in the pathogenesis of transmissible  
CC spongiform encephalopathies (TSEs). Examples of TSEs include Bovine  
CC (CJD) and Gerstmann-Strausler-Sheinker syndrome (GSS). The concentration  
CC of this protein in body fluid or tissue samples may be measured by an  
CC assay of the present invention, in which a PrP epitope is captured by an  
CC antibody, which is then detected. The presence of PrP indicates TSE. PrP  
CC epitopes (AAB07320-B07326) are derived from the protease resistant core  
CC of PrP that is occluded when the PrP is in an aggregated state.  
XX  
SQ Sequence 208 AA:  
  
Query Match 100.0%; Score 55; DB 21; Length 208;  
Best Local Similarity 100.0%; Pred. No. 0.43; Mismatches 0; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GWGDPHGG 8  
ID 41 GWGDPHGG 48  
DB  
  
RESULT 13  
AAB07318  
ID AAB07318 standard; protein; 208 AA.  
XX  
AC AAB07318;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Human prion protein sequence.  
XX  
DE Human prion protein; transmissible spongiform encephalopathy;  
KW bovine spongiform encephalopathy; BSE diagnosis; TSE; PrP.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FH Region 29..69  
FT /note="Repeat region consisting of tandem repeats  
FT of repeat unit: PHGGGWGQ (AAB07319)"  
FT Disulfide-bond 157..192  
FT Modified-site 208  
FT /note="C-terminal phospho-inositol glycolipid  
FT membrane anchor (-GPI)"  
XX  
XX WO200029850-A1.  
XX  
XX 25-MAY-2000.  
XX  
XX 27-OCT-1999; 99WO-FI00897.  
XX  
XX 17-NOV-1998; 98FI-0002481.  
XX  
XX (WALL-) WALLAC OY.  
XX (BBSR-) BBSRC OFFICE.  
XX  
XX Hope J, Barnard GJR, Birkett CR,  
XX  
XX WPI; 2000-387880/33.  
XX  
XX Novel immunoassay for prion protein, used for the determination of  
XX transmissible spongiform encephalopathies in bovines -  
XX  
XX Disclosure; Page 43-44; 50pp; English.

CC The present sequence is the human prion protein (PrP) sequence.  
CC Conversion of the normal cellular form of PrP into an aggregated,  
CC insoluble isoform is implicated in the pathogenesis of Transmissible  
CC Spongiform Encephalopathies (TSEs). Examples of TSEs include Bovine  
CC Spongiform Encephalopathy (BSE), scrapie, Creutzfeldt-Jakob disease  
CC (CJD) and Gerstmann-Strausler-Sheinker syndrome (GSS). The concentration  
CC of this protein in body fluid or tissue samples may be measured by an  
CC assay of the present invention, in which a PrP epitope is captured by an  
CC antibody, which is then detected. The presence of PrP indicates TSE. PrP  
CC epitopes (AAB07320-B07326) are derived from the protease resistant core  
CC of PrP that is occluded when the PrP is in an aggregated state.

CC Sequence 208 AA:

Query Match 100.0%; Score 55; DB 21; Length 208;  
Best Local Similarity 100.0%; Pred. No. 0.43;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGDPHGG 8  
|||||||  
34 GWGDPHGG 41

RESULT 14

AAB07327 standard; protein; 208 AA.

AC AAB07327;

DT 17-OCT-2000 (first entry)

DE Mouse prion protein sequence.

KW Mouse; prion protein; transmissible spongiform encephalopathy;  
KW bovine spongiform encephalopathy; TSE diagnosis; PrP.

OS Mus sp.

XX Location/Qualifiers

FT Key 37..68  
FT Region /note= "Repeat region consisting of tandem repeats  
FT of repeat unit: PHGGWGQ (AAB07319)"

FT Disulfide-bond 156..191

FT Modified-site 208

FT /note= "C-terminal phospho-inositol glycolipid  
FT membrane anchor (-GPI)"

XX WO200029849-A1.

XX 25-MAY-2000.

PF 27-OCT-1999; 99WO-FI00896.

PR 17-NOV-1998; 98FI-0002480.

PA (WALL-) WALLAC OY.  
PA (BBSR-) BBSRC OFFICE.

PI Hope J, Barnard GJR, Birkett CR;

XX WPI; 2000-399778/34.

PT New immunoassay for prion protein, used for determination of  
PT transmissible spongiform encephalopathies in mammals, comprises  
PT specific capture antibody -

PS Disclosure: Page 41-42; 50pp; English.

CC The present sequence is the mouse prion protein (PrP) sequence.

CC Conversion of the normal cellular form of PrP into an aggregated,  
CC insoluble isoform is implicated in the pathogenesis of Transmissible  
CC Spongiform Encephalopathies (TSEs). Examples of TSEs include Bovine  
CC Spongiform Encephalopathy (BSE), scrapie, Creutzfeldt-Jakob disease

CC (CJD) and Gerstmann-Strausler-Sheinker syndrome (GSS). The concentration  
CC of this protein in body fluid or tissue samples may be measured by an  
CC assay of the present invention, in which a PrP epitope is captured by an  
CC antibody, which is then detected. The presence of PrP indicates TSE. PrP  
CC epitopes (AAB07320-B07326) are derived from the protease resistant core  
CC of PrP that is occluded when the PrP is in an aggregated state.

CC Sequence 208 AA:

Query Match 100.0%; Score 55; DB 21; Length 208;  
Best Local Similarity 100.0%; Pred. No. 0.43;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGDPHGG 8  
|||||||  
DB 41 GWGDPHGG 48

RESULT 15

AAB07329 standard; protein; 208 AA.

AC AAB07329;

DT 17-OCT-2000 (first entry)

DE Human prion protein sequence.

KW Human; prion protein; transmissible spongiform encephalopathy;  
KW bovine spongiform encephalopathy; TSE diagnosis; PrP.

OS Homo sapiens.

XX Location/Qualifiers

FT Key 29..69  
FT Region /note= "Repeat region consisting of tandem repeats  
FT of repeat unit: PHGGWGQ (AAB07319)"

FT Disulfide-bond 157..192

FT Modified-site 208

FT /note= "C-terminal phospho-inositol glycolipid  
FT membrane anchor (-GPI)"

XX WO200029849-A1.

XX 25-MAY-2000.

PF 27-OCT-1999; 99WO-FI00896.

PR 17-NOV-1998; 98FI-0002480.

PA (WALL-) WALLAC OY.  
PA (BBSR-) BBSRC OFFICE.

PI Hope J, Barnard GJR, Birkett CR;

XX WPI; 2000-399778/34.

PT New immunoassay for prion protein, used for determination of  
PT transmissible spongiform encephalopathies in mammals, comprises  
PT specific capture antibody -

PS Disclosure: Page 43-44; 50pp; English.

CC The present sequence is the human prion protein (PrP) sequence.

CC Conversion of the normal cellular form of PrP into an aggregated,  
CC insoluble isoform is implicated in the pathogenesis of Transmissible  
CC Spongiform Encephalopathies (TSEs). Examples of TSEs include Bovine  
CC Spongiform Encephalopathy (BSE), scrapie, Creutzfeldt-Jakob disease

CC (CJD) and Gerstmann-Strausler-Sheinker syndrome (GSS). The concentration  
CC of this protein in body fluid or tissue samples may be measured by an  
CC assay of the present invention, in which a PrP epitope is captured by an

CC antibody, which is then detected. The presence of PrP indicates TSE. PrP  
CC epitopes (AAB07320-B07326) are derived from the protease resistant core

CC of Prp that is occluded when the Prp is in an aggregated state.

XX Sequence 208 AA;

Query Match 100.0%; Score 55; DB 21; Length 208;

Best Local Similarity 100.0%; Pred. No. 0.43;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GNGOPHGG 8

DB 34 GNGOPHGG 41

Search completed: January 3, 2003, 15:28:34  
Job time : 36.4545 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:00 ; Search time 13.2727 Seconds  
(without alignments)  
57.944 Million cell updates/sec

Title: US-09-543-188a-1  
Perfect score: 55  
Sequence: 1 GWGQPHHG 8

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
1 number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	226	2 A53892	prion-related prot
2	55	100.0	232	2 S71041	major prion protei
3	55	100.0	239	2 S53633	major prion protei
4	55	100.0	241	2 S71056	major prion protei
5	55	100.0	241	2 S71048	major prion protei
6	55	100.0	245	2 S53627	major prion protei
7	55	100.0	245	2 S71045	major prion protei
8	55	100.0	252	2 I61848	major prion protei
9	55	100.0	252	2 S53634	major prion protei
10	55	100.0	252	2 S53631	major prion protei
11	55	100.0	252	2 JC6175	major prion protei
12	55	100.0	253	1 UJHU	major prion protei
13	55	100.0	253	2 I37032	major prion protei
14	55	100.0	253	2 I61847	major prion protei
15	55	100.0	253	2 S53635	major prion protei
16	55	100.0	253	2 I84423	major prion protei
17	55	100.0	253	2 S53618	major prion protei
18	55	100.0	253	2 S53619	major prion protei
19	55	100.0	253	2 S53620	major prion protei
20	55	100.0	253	2 S71055	major prion protei
21	55	100.0	253	2 S53623	major prion protei
22	55	100.0	253	2 S53624	major prion protei
23	55	100.0	253	2 S53625	major prion protei
24	55	100.0	253	2 S53617	major prion protei
25	55	100.0	253	2 S53614	major prion protei
26	55	100.0	253	2 S53616	major prion protei
27	55	100.0	254	1 UJHYH	major prion prp-Sc
28	55	100.0	254	2 A34759	prion protein - Ch
29	55	100.0	254	2 B34759	prion protein - go

30	55	100.0	254	2 A23544	major prion protei
31	55	100.0	256	2 JU0268	major prion protei
32	55	100.0	256	2 S7149	major prion protei
33	55	100.0	256	2 A54281	major prion protei
34	55	100.0	257	2 A23545	major prion protei
35	55	100.0	257	2 JU1900	major prion protei
36	55	100.0	260	2 S53629	major prion protei
37	55	100.0	264	2 S37137	major prion protei
38	55	100.0	264	2 A54330	major prion protei
39	43	78.2	381	2 A44323	pentaxin prx3 prec
40	41	74.5	395	2 T08350	hypothetical prote
41	41	74.5	521	2 E64862	probable membrane
42	40	72.7	327	2 E87218	probable prephenat
43	40	72.7	1161	2 T45294	hypothetical prote
44	39	70.9	69	2 B89016	protein B0213.5 li
45	39	70.9	120	2 F86824	hypothetical prote

## ALIGNMENTS

RESULT 1  
A53892  
prion-related protein - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 13-Aug-1999  
C:Accession: A53892  
R:Liao, Y.C.; Tokes, Z.; Lim, E.; Lackey, A.; Woo, C.H.; Button, J.D.; Clawson, G.A.  
Lab. Invest. 57, 370-374, 1987  
A:Title: Cloning of rat "prion-related protein" cDNA.  
A:Reference number: A53892; MUID:88037055; PMID:2889848  
A:Accession: A53892  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-226 <LIA>  
A:Cross-references: GB:M20313; NID:g206391; PIDN:AAA1947.1; PID:g206392  
C:Superfamily: major prion protein

Query Match 100.0%; Score 55; DB 2; Length 226;  
Best Local Similarity 100.0%; Pred. No. 0.069;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 GWGQPHHG 8  
DB 36 GWGQPHHG 43  
|||||||

RESULT 2  
major prion protein - black-handed spider monkey (fragment)  
C:Species: Ateles geoffroyi (black-handed spider monkey)  
C:Date: 27-Oct-1996 #sequence\_revision 07-Feb-1997 #text\_change 13-Aug-1999  
R:Schatz, H.M.  
submitted to the EMBL Data Library, April 1994  
A:Reference number: S71041  
A:Accession: S71041  
A:Molecule type: DNA  
A:Residues: 1-232 <SCH>  
A:Cross-references: EMBL:U08309; NID:g474376; PIDN:AAC50097.1; PID:g474377  
R:Schatz, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.  
J. Mol. Biol. 245, 362-374, 1995  
A:Title: Prion protein gene variation among primates.  
A:Reference number: S53614; MUID:95139066; PMID:7837269  
A:Accession: S53630  
A>Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-194, 'R', 196-231 <SCW>  
A:Cross-references: EMBL:U08309  
C:Superfamily: major prion protein  
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match 100.0%; Score 55; DB 2; Length 232;





major prion protein - Cercopithecus diana  
C:Species: Cercopithecus diana  
C:Date: 14-Feb-1997 #sequence\_revision 14-Feb-1997 #text\_change 13-Aug-1999  
C:Accession: S71045; S53628  
R:Schaetzl, H.M.  
submitted to the EMBL Data Library, April 1994  
A:Reference number: S71041  
A:Accession: S71045  
A:Molecule type: DNA  
A:Residues: 1-245 <SCH>  
A:Cross-references: EMBL:U08292; NID:q474342; PIDN:AAC50081.1; PID:q474343  
R:Schaetzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.  
J. Mol. Biol. 245, 362-374, 1995  
A:Title: Prion protein gene variation among primates.  
A:Reference number: S53614; MUID:95139066; PMID:7837269  
A:Accession: S53628  
A:Molecule type: DNA  
A:Status: nucleic acid sequence not shown  
A:Residues: 8-10,'L',12-202,'R',204-239 <SCW>  
A:Cross-references: EMBL:U08292  
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane protein

Query Match 100.0%; Score 55; DB 2; Length 245;  
Best Local Similarity 100.0%; Pred. No. 0.074;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8  
|||||  
Db 56 GWGQPHGG 63

RESULT 8  
161848  
major prion protein precursor - common squirrel monkey  
C:Species: Saimiri sciureus (common squirrel monkey)  
C:Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 13-Aug-1999  
C:Accession: I61848  
R:Cervenakova, L.; Brown, P.; Goldfarb, L.G.; Nagle, J.; Pettrone, K.; Rubenstein, R.; D  
Proc. Natl. Acad. Sci. U.S.A. 91, 12159-12162, 1994  
A:Title: Infectious amyloid precursor gene sequences in primates used for experimental t  
A:Reference number: I36907; MUID:95083661; PMID:7991600  
A:Accession: I61848  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-252 <RES>  
A:Cross-references: EMBL:U15165; NID:9595852; PIDN:AAA68636.1; PID:9595853  
C:Superfamily: major prion protein

Query Match 100.0%; Score 55; DB 2; Length 252;  
Best Local Similarity 100.0%; Pred. No. 0.076;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8  
|||||  
Db 55 GWGQPHGG 62

RESULT 9  
553634  
major prion protein - common marmoset  
C:Species: Callithrix jacchus (common marmoset)  
C:Date: 28-Oct-1996 #sequence\_revision 07-Feb-1997 #text\_change 13-Aug-1999  
C:Accession: S53634; S71047  
R:Schaetzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.  
J. Mol. Biol. 245, 362-374, 1995  
A:Title: Prion protein gene variation among primates.  
A:Reference number: S53614; MUID:95139066; PMID:7837269  
A:Accession: S53634  
A:Molecule type: DNA  
A:Status: nucleic acid sequence not shown  
A:Residues: 1-252 <SCH>  
A:Cross-references: EMBL:U08304

R:Schaetzl, H.M.  
submitted to the EMBL Data Library, April 1994  
A:Reference number: S71041  
A:Accession: S71047  
A:Molecule type: DNA  
A:Residues: 1-209,'E',211-252 <SCW>  
A:Cross-references: EMBL:U08304; NID:q474366; PIDN:AAC50092.1; PID:q474367  
C:Superfamily: major prion protein  
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match 100.0%; Score 55; DB 2; Length 252;  
Best Local Similarity 100.0%; Pred. No. 0.076;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8  
|||||  
Db 55 GWGQPHGG 62

RESULT 10  
553631  
major prion protein - brown capuchin  
C:Species: Cebus apella (brown capuchin, black-capped capuchin)  
C:Date: 28-Oct-1996 #sequence\_revision 07-Feb-1997 #text\_change 13-Aug-1999  
C:Accession: S53631; S71044  
R:Schaetzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.  
J. Mol. Biol. 245, 362-374, 1995  
A:Title: Prion protein gene variation among primates.  
A:Reference number: S53614; MUID:95139066; PMID:7837269  
A:Accession: S53631  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-252 <SCH>  
A:Cross-references: EMBL:U08295  
R:Schaetzl, H.M.  
submitted to the EMBL Data Library, April 1994  
A:Reference number: S71041  
A:Accession: S71044  
A:Molecule type: DNA  
A:Residues: 1-209,'E',211-252 <SCW>  
A:Cross-references: EMBL:U08295; NID:q474348; PIDN:AAC50084.1; PID:q474349  
C:Superfamily: major prion protein  
C:Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane

Query Match 100.0%; Score 55; DB 2; Length 252;  
Best Local Similarity 100.0%; Pred. No. 0.076;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8  
|||||  
Db 55 GWGQPHGG 62

RESULT 11  
J66175  
prion protein - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 11-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 13-Aug-1999  
C:Accession: J66175  
R:Loftus, B.; Rogers, M.  
Gene 184, 215-219, 1997  
A:Title: Characterization of a prion protein (PrP) gene from rabbit: a species with a  
A:Reference number: J66175; MUID:97183665; PMID:9031631  
A:Accession: J66175  
A:Molecule type: DNA  
A:Residues: 1-252 <LOF>  
A:Cross-references: GB:U08334; NID:q1490412; PIDN:AAC48697.1; PID:q1490413  
C:Comment: This protein is a cellular protein, it is involved in the neurodegenerativ  
A:Gene: PrP  
A:Superfamily: major prion protein  
C:Keywords: disulfide bond; prion

```

Query Match          100.0%; Score 55; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GWGQPHGG 8
DB      56 GWGQPHGG 63

RESULT 12
UTRU
major prion protein precursor - human
N:Alternate names: 11k amyloid protein; 27-30k sialoglycoprotein; PrP 27-30; PrP 33-35C;
C:Species: Homo sapiens (man)
C:Date: 25-Oct-1987 #sequence_revision 12-Apr-1996 #text_change 16-Jun-2000
C:Accession: A24173; A40372; A05017; S14078; I54322; I68597; I58135; I59184; I79633; I79
R:Kretzschmar, H.A.; Stowring, L.E.; Westaway, D.; Studdiblane, W.H.; Prusiner, S.B.; De
DNA 5, 315-324, 1986
A:Title: Molecular cloning of a human prion protein CDNA.
A:Reference number: A24173; MUID:86300093; PMID:3755672
A:Accession: A24173
A:Molecule type: mRNA
A:Residues: 1-253 <KRE>
A:Cross-references: GB:M13899; NID:9190467; PIDN:AAA60182.1; PID:9190468
R:Puckett, C.; Concanon, P.; Casey, C.; Hood, L.
Am. J. Hum. Genet. 49, 320-329, 1991
A:Title: Genomic structure of the human prion protein gene.
A:Reference number: A40372; MUID:91328137; PMID:1678248
A:Accession: A40372
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-80, 89-253 <PUC>
A:Cross-references: GB:X83416; NID:9747846; PIDN:CAA58442.1; PID:9747847
A:Note: the deletion may be a polymorphism; the alternative deletion of 82-89 could not
R:Liso, Y.C.J.; Lebo, R.V.; Clawson, G.A.; Smuckler, E.A.
Science 233, 364-367, 1986
A:Reference number: A05017; MUID:86261778; PMID:3014653
A:Accession: A05017
A:Molecule type: mRNA
A:Residues: 8-117, 119-253 <LIA>
A:Cross-references: GB:D00015; NID:9220015; PIDN:BAA0011.1; PID:9220016; GB:M13667; NID
R:Tagliavini, F.; Prelli, F.; Ghiso, J.; Bugiani, O.; Serban, D.; Prusiner, S.B.; Farlow
EMBO J. 10, 513-519, 1991
A:Title: Amyloid protein of Gerstmann-Straussler-Scheinker disease (Indiana kindred) is
A:Reference number: S14078; MUID:91160504; PMID:1672107
A:Accession: S14078
A:Molecule type: protein
A:Residues: 58-72, 'X', 74-76, 'XX', 79, 'XXX', 83-86, 111-128, 'V', 130-150 <TAG>
R:Gierlich, J.F.; Knopman, D.S.; List, J.F.; Olson, K.; Frey, W.H.
Mol. Genet. 1, 443-444, 1992
A:Title: Deletion in the prion protein gene in a demented patient.
A:Reference number: I54322; MUID:93250789; PMID:1363802
A:Accession: I54322
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 9-83, 92-240 <RES>
A:Cross-references: GB:M81929; NID:9190517; PIDN:AAB59442.1; PID:9190518
A:Accession: I68597
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 8-240 <RES>
A:Cross-references: GB:M81930; NID:9190519; PIDN:AAB59443.1; PID:9190520
R:Brown, P.; Goldfarb, L.G.; McComble, W.R.; Nieto, A.; Squillacote, D.; Sheremata, W.;
Neurology 42, 422-427, 1992
A:Title: Atypical Creutzfeldt-Jakob disease in an American family with an insert mutatio
A:Reference number: I58135; MUID:92140671; PMID:1736177
A:Accession: I58135
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 51-91, 'PHGGGCGPRHGGGCGPRHGGGCGPRHGGGCGPRHGGGCG' <RE2>
A:Cross-references: GB:S80539; NID:9244698; PIDN:AAB21334.1; PID:9244699
R:Goldfarb, L.G.; Brown, P.; McComble, W.R.; Goldfarb, D.; Swergold, G.D.; Willis, P.R.;
Proc. Natl. Acad. Sci. U.S.A. 88, 10926-10930, 1991

```

```

A:Title: Transmissible familial Creutzfeldt-Jakob disease associated with five, seven
A:Reference number: I59184; MUID:92073400; PMID:1683708
A:Accession: I59184
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 60-67 <GOL>
A:Cross-references: GB:S71208; NID:9239877; PIDN:AAB20521.1; PID:9239878; GB:S71210;
C:Genetics:
A:Gene: GDB:PRNP; CJD; PrP
A:Cross-references: GDB:120720; OMIM:176640; OMIM:137440
A:Map position: 20pter-20p12
A:Introns: #status absent
A:Note: one intron occurs before the initiator codon
C:Superfamily: this gene is associated with Creutzfeldt-Jakob disease (CJD), Gerstmann-Strau
C:Keywords: amyloid; blocked carboxyl end; brain; glycoprotein; lipoprotein; phosphat
F:11-22/Domain: signal sequence #status predicted <SIG>
F:23-230/Product: major prion protein #status predicted <MAT>
F:54-92/Region: 8-residue repeats (P-H-G-G-W-G-Q)
F:112-134/Domain: transmembrane #status predicted <TM1>
F:231-253/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:179-214/Disulfide bonds: #status predicted
F:181,197/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:230/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature f
Query Match          100.0%; Score 55; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GWGQPHGG 8
DB      56 GWGQPHGG 63

RESULT 13
major prion protein precursor - gorilla
C:Species: Gorilla gorilla (gorilla)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 13-Aug-1999
C:Accession: I37032
R:Cervenaikova, L.; Brown, P.; Goldfarb, L.G.; Nagle, J.; Pettrone, K.; Rubenstein, R.
Proc. Natl. Acad. Sci. U.S.A. 91, 12159-12162, 1994
A:Title: Infectious amyloid precursor gene sequences in primates used for experimenta
A:Reference number: I36907; MUID:95083661; PMID:7991600
A:Accession: I37032
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-253 <RES>
A:Cross-references: EMBL:U15166; NID:9563208; PIDN:AAA68633.1; PID:9563209
C:Superfamily: major prion protein_
Query Match          100.0%; Score 55; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GWGQPHGG 8
DB      56 GWGQPHGG 63

RESULT 14
major prion protein precursor - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 13-Aug-1999
C:Accession: I61847; S71060; S53615
R:Cervenaikova, L.; Brown, P.; Goldfarb, L.G.; Nagle, J.; Pettrone, K.; Rubenstein, R.
Proc. Natl. Acad. Sci. U.S.A. 91, 12159-12162, 1994
A:Title: Infectious amyloid precursor gene sequences in primates used for experimenta
A:Reference number: I36907; MUID:95083661; PMID:7991600
A:Accession: I61847
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA

```

A:Residues: 1-253 <RES>  
 A:Cross-references: EMBL:U15039; NID:g609303; PIDN:AAA68632.1; PID:g609304  
 R:Schatzl, H.M.  
 submitted to the EMBL Data Library, April 1994  
 A:Reference number: S71041  
 A:Accession: S71060  
 A:Molecule type: DNA  
 A:Residues: 1-253 <SCW>  
 A:Cross-references: EMBL:U08296; NID:g474350; PIDN:AAC50085.1; PID:g474351  
 R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.  
 J. Mol. Biol. 245, 362-374, 1995  
 A:Title: Prion protein gene variation among primates.  
 A:Reference number: S53614; MUID:95139066; PMID:7837269  
 A:Accession: S53615  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-210, 'R', 212-253 <SCH>  
 A:Cross-references: EMBL:U08296  
 C:Superfamily: major prion protein  
 Keywords: amyloid; brain; glycoprotein; lipoprotein; prion; scrapie; transmembrane protein

Query Match 100.0%; Score 55; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 0.076;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNGOPHGG 8  
 |||||||  
 Db 56 GNGOPHGG 63

RESULT 15  
 S53635  
 prion protein - simang  
 C:Species: Hylobates syndactylus (simang)  
 C:Date: 15-Jul-1995 #sequence\_revision 19-Apr-1996 #text\_change 13-Aug-1999  
 C:Accession: S53635  
 R:Schatzl, H.M.; da Costa, M.; Taylor, L.; Cohen, F.E.; Prusiner, S.B.  
 J. Mol. Biol. 245, 362-374, 1995  
 A:Title: Prion protein gene variation among primates.  
 A:Reference number: S53614; MUID:95139066; PMID:7837269  
 A:Accession: S53635  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-253 <SCH>  
 A:Cross-references: EMBL:U08308; NID:g474374; PIDN:AAC50096.1; PID:g474375  
 A:Note: the source was designated as Symphalangus syndactylus  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994  
 Superfamily: major prion protein

Query Match 100.0%; Score 55; DB 2; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 0.076;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GNGOPHGG 8  
 |||||||  
 Db 56 GNGOPHGG 63

Search completed: January 3, 2003, 15:33:13  
 Job time : 13.2727 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:22:10 : Search time 6.90809 Seconds  
(without alignments)  
48.025 Million cell updates/sec

Title: US-09-543-188a-1  
Perfect score: 55  
Sequence: 1 GWCOPHC8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	55	100.0	232	1	PRIO_ATEGE
2	55	100.0	238	1	PRIO_CERAT
3	55	100.0	238	1	PRIO_CERAT
4	55	100.0	239	1	PRIO_AOTTR
5	55	100.0	241	1	PRIO_CAMMO
6	55	100.0	241	1	PRIO_MANSF
7	55	100.0	245	1	PRIO_CERAE
8	55	100.0	246	1	PRIO_CERMA
9	55	100.0	246	1	PRIO_CERMA
10	55	100.0	246	1	PRIO_CERMA
11	55	100.0	252	1	PRIO_ATEPA
12	55	100.0	252	1	PRIO_CALUA
13	55	100.0	252	1	PRIO_CEBAP
14	55	100.0	252	1	PRIO_RABIT
15	55	100.0	253	1	PRIO_COLGU
16	55	100.0	253	1	PRIO_GORGO
17	55	100.0	253	1	PRIO_HUMAN
18	55	100.0	253	1	PRIO_MACRA
19	55	100.0	253	1	PRIO_PANTR
20	55	100.0	253	1	PRIO_PONPY
21	55	100.0	253	1	PRIO_PREFR
22	55	100.0	254	1	PRIO_CRIGR
23	55	100.0	254	1	PRIO_CRIMI
24	55	100.0	254	1	PRIO_MESAU
25	55	100.0	254	1	PRIO_MOUSE
26	55	100.0	254	1	PRIO_RAT
27	55	100.0	254	1	PRIO_SIGHI
28	55	100.0	255	1	PRIO_CAMDR
29	55	100.0	255	1	PRIO_CANFA
30	55	100.0	256	1	PRIO_CAPII
31	55	100.0	256	1	PRIO_CEREL
32	55	100.0	256	1	PRIO_FELICA
33	55	100.0	256	1	PRIO_ODOHE

34	55	100.0	256	1	PRIO_SHEEP	P23907 ovis aries
35	55	100.0	256	1	PRP2_BOVIN	O01880 bos taurus
36	55	100.0	256	1	PRP2_TRAST	P40243 tragelaphus
37	55	100.0	257	1	PRIO_MOUSEP	P52114 mustela put
38	55	100.0	257	1	PRIO_MUSVI	P40244 mustela vis
39	55	100.0	257	1	PRIO_PIG	P49927 sus scrofa
40	55	100.0	260	1	PRIO_SAISC	P40258 saimiri sci
41	55	100.0	264	1	PRIO_BOVIN	P10279 bos taurus
42	55	100.0	264	1	PRP1_TRAST	P40242 tragelaphus
43	49	89.1	259	1	PRIO_TRIUV	P51780 trichosurus
44	43	78.2	381	1	PRX3_HUMAN	P26022 homo sapien
45	41	74.5	507	1	YCGG_ECOLI	P75995 escherichia

## ALIGNMENTS

RESULT 1  
PRIO\_ATEGE STANDARD: PRT: 232 AA.  
AC P40246: 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Major prion protein precursor (Prp) (PrP27-30) (PrP33-35C) (Fragment).  
GN PRNP.  
OS Ateles geoffroyi (Black-handed spider monkey).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.  
OX NCBI\_TaxID=9509;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95139066; PubMed=7837269;  
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;  
RT "Prion protein gene variation among primates."  
RL J. Mol. Biol. 245:362-374(1995).  
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE  
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.  
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED  
CC "RODS".  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -!- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND  
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,  
CC CRUZFELDT-JAKOB DISEASE (CJD), GERSTMAN-STRAUSSLER SYNDROME  
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),  
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.  
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: U08309; AAC50097.1; -.  
CC HSSP: P04156; IEIG.  
CC InterPro: IPR00817; Prion.  
CC Pfam: PF00377; prion.1.  
CC SMART: SM00157; PRP.1.  
CC PROSITE: PS00291; PRION\_1; 1.  
CC PROSITE: PS00706; PRION\_2; 1.  
CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.  
CC KW NON\_TER  
CC FT SIGNAL 1 15  
CC FT CHAIN 16 214  
CC FT PROPEP 215 >232  
CC FT LIPID 214 214  
CC FT DISULFID 163 198  
CC FT CARBOHYD 165 165  
CC FT CARBOHYD 181 181  
CC FT DOMAIN 44 84  
CC 4 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-

FT REPEAT 44 51 0.  
FT REPEAT 52 59 1.  
FT REPEAT 60 67 2.  
FT REPEAT 68 75 3.  
FT REPEAT 232 232 4.  
SQ SEQUENCE 232 AA: 25596 MW: 0E2D75F04C05CC4A CRC64;

Query Match 100.0%; Score 55; DB 1; Length 232;  
Best Local Similarity 100.0%; Pred. No. 0.04;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWCOPHGG 8  
|||||||  
Db 48 GWCOPHGG 55

## RESULT 2

PRIOT.CERAT STANDARD; PRT; 238 AA.  
095145: 095200:  
01-NOV-1997 (Rel. 35, Created)  
01-NOV-1997 (Rel. 35, Last sequence update)  
15-JUN-2002 (Rel. 41, Last annotation update)  
DE Major prion protein precursor (Prp) (Prp27-30) (Prp33-35C) (Fragment).  
GN PRNP.  
OS Cercopithecus aethiops, and  
OS Macaca sylvanus (Barbary ape).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Cercopithecus.  
OX NCBI\_TaxID=36222, 9546;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA der Kuyt A.C., Dekker J.T., Goudsmit J.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE  
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.  
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED  
CC "RODS".  
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND  
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,  
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME  
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),  
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.  
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
CC EMBL: U75384; AAB50623.1; -  
CC EMBL: U75382; AAB50629.1; -  
CC HSSP: P04925; IAG2.  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; prion.1.  
DR SMART: SM00157; PRP.1.  
DR PROSITE: PS00291; PRION\_1; 1.  
DR PROSITE: PS00706; PRION\_2; 1.  
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 15  
FT CHAIN 16 215 BY SIMILARITY.  
FT PROPEP 216 238 MAJOR PRION PROTEIN.  
FT LIPID 215 215 REMOVED IN MATURE FORM (BY SIMILARITY).  
FT DISULFID 164 199 BY SIMILARITY.  
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT DOMAIN 44 76 4 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-  
FT REPEAT 44 52 0.  
FT REPEAT 53 60 1.  
FT REPEAT 61 68 2.  
FT REPEAT 69 76 3.  
SQ SEQUENCE 238 AA: 26123 MW: 5F59A3EBC3E3531B CRC64;

Query Match 100.0%; Score 55; DB 1; Length 238;  
Best Local Similarity 100.0%; Pred. No. 0.041;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWCOPHGG 8  
|||||||  
Db 49 GWCOPHGG 56

## RESULT 3

PRIOT.THEGE STANDARD; PRT; 238 AA.  
095270:  
01-NOV-1997 (Rel. 35, Created)  
01-NOV-1997 (Rel. 35, Last sequence update)  
01-NOV-1997 (Rel. 35, Last annotation update)  
DE Major prion protein precursor (Prp) (Prp27-30) (Prp33-35C) (Fragment).  
GN PRNP OR PRP.  
OS Theropithecus gelada (Gelada baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Theropithecus.  
OX NCBI\_TaxID=9565;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA der Kuyt A.C., Dekker J.T., Goudsmit J.;  
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE  
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.  
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED  
CC "RODS".  
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND  
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,  
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME  
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),  
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.  
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
CC EMBL: U75383; AAB50630.1; -  
CC HSSP: P04925; IAG2.  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; prion.1.  
DR SMART: SM00157; PRP.1.  
DR PROSITE: PS00291; PRION\_1; 1.  
DR PROSITE: PS00706; PRION\_2; 1.  
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 15  
FT CHAIN 16 215 BY SIMILARITY.  
FT DISULFID 164 199 MAJOR PRION PROTEIN.  
FT CARBOHYD 166 166 BY SIMILARITY.  
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DOMAIN 44 83 4 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-  
FT REPEAT 44 52 0.  
FT REPEAT 53 60 1.  
FT REPEAT 61 68 2.

FT REPEAT 61 68 3.  
FT REPEAT 69 76 4.  
FT NON\_TER 238 238  
SQ SEQUENCE 238 AA; 26104 MW; 5F59BFF602243HDB CRC64;

Query Match  
Best Local Similarity 100.0%; Score 55; DB 1; Length 238;  
Matches 8; Conservative 0; Pred. No. 0.041; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGPHG 8  
Db 49 GWGPHG 56

## RESULT 4

PRIO\_AOTTR STANDARD: PRT; 239 AA.  
AC P40245;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).  
GN PRNP.  
OS Aotus trivirgatus (Night monkey) (Douroucoulli).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.  
OX NCBI\_TaxID=9505;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95139066; PubMed=7837269;  
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;  
RT "Prion protein gene variation among primates."  
RL J. Mol. Biol. 245:362-374(1995).

-1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE  
HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.  
-1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED  
"RODS".

-1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
-1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND  
ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,  
CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME  
(GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),  
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.  
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U08293; AAC50082.1; -.  
CC HSSP: P04925; IAG2.  
CC InterPro: IPR000817; Prion.  
CC Pfam: PF00377; prion.1.  
CC SMART: SM00157; PRP.1.

DR PROSITE: PS00291; PRION\_1; 1.  
DR PROSITE: PS00706; PRION\_2; 1.  
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 15  
FT CHAIN 16 >239  
FT DISULFID 171 206  
FT CARBOHYD 173 173  
FT CARBOHYD 189 189  
FT DOMAIN 44 83

FT REPEAT 44 51  
FT REPEAT 52 59  
FT REPEAT 60 67  
FT REPEAT 68 75  
FT REPEAT 4. 4.  
FT REPEAT 1. 1.  
FT REPEAT 2. 2.  
FT REPEAT 3. 3.  
FT REPEAT 4. 4.

FT REPEAT 76 83 5.  
FT NON\_TER 239 239  
SQ SEQUENCE 239 AA; 26246 MW; 2EFB77E354B7024A CRC64;

Query Match  
Best Local Similarity 100.0%; Score 55; DB 1; Length 239;  
Matches 8; Conservative 0; Pred. No. 0.041; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGPHG 8  
Db 48 GWGPHG 55

## RESULT 5

PRIO\_CALMO STANDARD: PRT; 241 AA.  
AC P40248;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).  
GN PRNP.  
OS Callicebus moloch (Dusky titi).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Callcibinae;  
OC Callicebus.  
OX NCBI\_TaxID=9523;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95139066; PubMed=7837269;  
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;  
RT "Prion protein gene variation among primates."  
RL J. Mol. Biol. 245:362-374(1995).

-1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE  
HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.  
-1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED  
"RODS".

-1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
-1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND  
ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,  
CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME  
(GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),  
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.  
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U08312; AAC50100.1; -.  
CC HSSP: P04925; IAG2.  
CC InterPro: IPR000817; Prion.  
CC Pfam: PF00377; prion.1.  
CC SMART: SM00157; PRP.1.

DR PROSITE: PS00291; PRION\_1; 1.  
DR PROSITE: PS00706; PRION\_2; 1.  
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 15  
FT CHAIN 16 >241  
FT DISULFID 172 207  
FT CARBOHYD 174 174  
FT CARBOHYD 190 190  
FT DOMAIN 44 84

FT REPEAT 44 52  
FT REPEAT 53 60  
FT REPEAT 61 68  
FT REPEAT 69 76  
FT REPEAT 4. 4.  
FT REPEAT 1. 1.  
FT REPEAT 2. 2.  
FT REPEAT 3. 3.  
FT REPEAT 4. 4.

```
FT REPEAT 77 84 5.
FT NON_TER 241 241
SQ SEQUENCE 241 AA: 26373 MW: C6D2013EE7CAEC93 CRC64:

Query Match
Best Local Similarity 100.0%; Score 55; DB 1; Length 241;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWCOPHGC 8
Db 57 GWCOPHGC 64

RESULT 6
PRIO_MANSNP
ID PRIO_MANSNP STANDARD; PRT; 241 AA.
AC P40255;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C) (Fragment).
PRNP.
Mandillius sphinx (Mandril) (Papio sphinx).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Mandillius.
OX NCBI_Taxid=9561;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95139066; PubMed=7837269;
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
"Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "ROOS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U08303; AAC50091.1; -.
CC HSSP: P04925; IAG2.
DR InterPro: IPR000817; Prion.
DR Pfam: PF00377; prion; 1.
DR SMART: SM00157; prp; 1.
DR PROSITE: PS00291; PRION_1; 1.
DR PROSITE: PS00706; PRION_2; 1.
KW prion; Brain; Glycoprotein; GPI-anchor; Repeat; signal.
FT NON_TER 1 1
FT SIGNAL <1 15
FT CHAIN 16 223
FT PROPEP 224 >241
FT LIPID 223 223
FT DISULFID 172 207
FT CARBOHYD 174 174
FT CARBOHYD 190 190
FT DOMAIN 44 84
FT REPEAT 44 52
FT REPEAT 53 60
FT 2.
```

```
FT REPEAT 61 68 3.
FT REPEAT 69 76 4.
FT REPEAT 77 84 5.
FT NON_TER 241 241
SQ SEQUENCE 241 AA: 26398 MW: E539D84E2E2B59DE CRC64:

Query Match
Best Local Similarity 100.0%; Score 55; DB 1; Length 241;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWCOPHGC 8
Db 49 GWCOPHGC 56

RESULT 7
PRIO_CERAE
ID PRIO_CERAE STANDARD; PRT; 245 AA.
AC P40250;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
DE PRNP.
OC Cercopithecus aethiops (Green monkey) (Grivet), and
OC Cercopithecus diana (Diana monkey).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_Taxid=9534, 36224;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95139066; PubMed=7837269;
RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
"Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "ROOS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U08291; AAC50080.1; -.
CC HSSP: P04925; IAG2.
DR InterPro: IPR000817; Prion.
DR Pfam: PF00377; prion; 1.
DR PRINTS: PR00341; PRION.
DR SMART: SM00157; prp; 1.
DR PROSITE: PS00291; PRION_1; 1.
DR PROSITE: PS00706; PRION_2; 1.
KW prion; Brain; Glycoprotein; GPI-anchor; Repeat; signal.
FT SIGNAL 1 22
FT CHAIN 23 222
FT PROPEP 223 222
FT LIPID 222 222
FT DISULFID 171 206
FT CARBOHYD 173 173
FT CARBOHYD 189 189
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
```



FT DOMAIN 51 83 4 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-  
FT REPEAT 51 59 0.  
FT REPEAT 60 67 1.  
FT REPEAT 68 75 2.  
FT REPEAT 76 83 3.  
FT REPEAT 76 83 4.  
SQ SEQUENCE 245 AA; D582B58E2726C99A CRC64;  
Query Match 100.0%; Score 55; DB 1; Length 245;  
Best Local Similarity 100.0%; Pred. No. 0.042;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GWGPHGG 8  
DB 56 GWGPHGG 63  
RESULT 8  
PRIO\_CERMO STANDARD; PRT; 246 AA.  
095172; 095173;  
01-NOV-1997 (Rel. 35, Created)  
01-NOV-1997 (Rel. 35, Last sequence update)  
15-00U-1998 (Rel. 36, Last annotation update)  
DE Major prion protein precursor (Prp) (Prp27-30) (Prp33-35C) (Fragment).  
GN PRNP.  
OS Cercopithecus mona, and  
OC Cercopithecus neglectus.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Cercopithecus.  
OX NCBI\_TaxID=36226, 36227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA der Kuyt A.C., Dekker J.T., Goudsmit J.;  
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE  
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.  
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED  
CC "RODS".  
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND  
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,  
CC CRETZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME  
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),  
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.  
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
DR EMBL: U75386; AAB50625.1; -  
DR EMBL: U75387; AAB50626.1; -  
DR HSSP: P04925; IAG2.  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; prion; 1.  
DR SMART: SM00157; PRP; 1.  
DR PROSITE: PS00291; PRION\_1; 1.  
DR PROSITE: PS00706; PRION\_2; 1.  
KW prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 15  
FT CHAIN 16 223 MAJOR PRION PROTEIN.  
FT PROPEP 224 246 REMOVED IN MATURE FORM (BY SIMILARITY).  
FT LIPID 223 223 GPI-ANCHOR (BY SIMILARITY).  
FT DISULFID 172 207 BY SIMILARITY.  
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT DOMAIN 44 84 5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-  
FT REPEAT 44 52 0.  
FT REPEAT 53 60 1.  
FT REPEAT 61 68 2.  
FT REPEAT 69 76 3.  
FT REPEAT 77 84 4.  
FT REPEAT 77 84 5.  
SQ SEQUENCE 246 AA; 26900 MW; 835D147CA2B4FDD3 CRC64;  
Query Match 100.0%; Score 55; DB 1; Length 246;  
Best Local Similarity 100.0%; Pred. No. 0.042;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GWGPHGG 8  
DB 49 GWGPHGG 56  
RESULT 9  
PRIO\_CERPA STANDARD; PRT; 246 AA.  
095174;  
01-NOV-1997 (Rel. 35, Created)  
01-NOV-1997 (Rel. 35, Last sequence update)  
01-NOV-1997 (Rel. 35, Last annotation update)  
DE Major prion protein precursor (Prp) (Prp27-30) (Prp33-35C) (Fragment).  
GN PRNP.  
OS Cercopithecus patas.  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Cercopithecus.  
OX NCBI\_TaxID=27677;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA der Kuyt A.C., Dekker J.T., Goudsmit J.;  
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE  
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.  
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED  
CC "RODS".  
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND  
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,  
CC CRETZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME  
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),  
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.  
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
DR EMBL: U75388; AAB50627.1; -  
DR HSSP: P04925; IAG2.  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; prion; 1.  
DR SMART: SM00157; PRP; 1.  
DR PROSITE: PS00291; PRION\_1; 1.  
DR PROSITE: PS00706; PRION\_2; 1.  
KW prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.  
FT NON\_TER 1 1  
FT SIGNAL <1 15  
FT CHAIN 16 223 MAJOR PRION PROTEIN.  
FT PROPEP 224 246 REMOVED IN MATURE FORM (BY SIMILARITY).  
FT LIPID 223 223 GPI-ANCHOR (BY SIMILARITY).  
FT DISULFID 172 207 BY SIMILARITY.  
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DOMAIN 44 84 5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-

```
FT REPEAT 44 52 0.
FT REPEAT 1. 53 60 1.
FT REPEAT 61 68 2.
FT REPEAT 69 76 3.
FT REPEAT 77 84 4.
FT REPEAT 84 84 5.
SQ SEQUENCE 246 AA: 26886 MW: D35D105BEC53108 CRC64:
Query Match 100.0%; Score 55; DB 1; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
Db 49 GWGQPHGG 56

RESULT 10
PRIO_CERTO STANDARD: PRT: 246 AA.
095176:
01-NOV-1997 (rel. 35, Created)
01-NOV-1997 (rel. 35, Last sequence update)
01-NOV-1997 (rel. 35, Last annotation update)
Major prion protein precursor (Prp) (Prp27-30) (Prp33-35C) (Fragment).
PRNP.
Cercopithecus torquatus atys (Red-crowned mangabey) (sooty mangabey).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9531;
RN [1]
RP SEQUENCE FROM N.A.
RA der Kuyi A.C., Dekker J.T., Goudsmit J.;
Submitted (Nov-1996) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CRETZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U75385; AAB50628.1; -.
CC DR HSSP: P04925; IAC2.
CC DR InterPro: IPR000817; Prion.
CC DR Pfam: PF00377; prion.1.
CC DR SMART: SM00157; prp.1.
CC DR PROSITE: PS00291; PRION_1; 1.
CC DR PROSITE: PS00706; PRION_2; 1.
CC KW prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
CC FT NON_TER 1 1
CC FT SIGNAL <1 15
CC FT CHAIN 16 223
CC FT PROPEP 224 246
CC FT LIPID 223 223
CC FT DISULFID 172 207
CC FT CARBOHYD 174 174
CC FT CAROHD 150 190
CC FT DOMAIN 44 84
CC Q.
BY SIMILARITY.
MAJOR PRION PROTEIN.
REMOVED IN MATURE FORM (BY SIMILARITY).
GPI-ANCHOR (BY SIMILARITY).
BY SIMILARITY.
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
Q.
```

```
FT REPEAT 44 52 1.
FT REPEAT 53 60 2.
FT REPEAT 61 68 3.
FT REPEAT 69 76 4.
FT REPEAT 77 84 5.
SQ SEQUENCE 246 AA: 26914 MW: F58679CBBC5ADC7 CRC64:
Query Match 100.0%; Score 55; DB 1; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8
Db 49 GWGQPHGG 56

RESULT 11
PRIO_ATEPA STANDARD: PRT: 252 AA.
P1446:
01-OCT-1996 (rel. 34, Created)
01-OCT-1996 (rel. 34, Last sequence update)
01-OCT-1996 (rel. 34, Last annotation update)
Major prion protein precursor (Prp) (Prp27-30) (Prp33-35C).
PRNP.
Ateles paniscus (Black spider monkey).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ateles.
OX NCBI_TaxID=9510;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95083661; PubMed=7991600;
RA Cervenakova L., Brown P., Goldfarb L.G., Nagle J., Petrone K.,
RA Rubenstein R., Dubnick M., Gibbs C.J., Gajdusek D.C.;
RT "Infectious amyloid precursor gene sequences in primates used for
RT experimental transmission of human spongiform encephalopathy.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:12159-12162(1994).
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CRETZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U15164; AAB68634.1; -.
CC DR HSSP: P04156; IE1G.
CC DR InterPro: IPR000817; Prion.
CC DR Pfam: PF00377; prion.1.
CC DR PRINTS: PR00341; PRION.
CC DR SMART: SM00157; prp.1.
CC DR PROSITE: PS00291; PRION_1; 1.
CC DR PROSITE: PS00706; PRION_2; 1.
CC KW prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
CC FT SIGNAL 1 22
CC FT CHAIN 23 229
CC FT PROPEP 230 252
CC FT LIPID 229 229
CC FT DISULFID 178 213
CC FT CARBOHYD 180 180
CC Q.
BY SIMILARITY.
MAJOR PRION PROTEIN.
REMOVED IN MATURE FORM (BY SIMILARITY).
GPI-ANCHOR (BY SIMILARITY).
BY SIMILARITY.
N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 51 90 5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
FT REPEAT 51 58 1.
FT REPEAT 59 66 2.
FT REPEAT 67 74 3.
FT REPEAT 75 82 4.
FT REPEAT 83 90 5.
SQ SEQUENCE 252 AA; 27718 MW; 20EA38A42DC56D1 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 252;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGPHGG 8
Db 55 GWGPHGG 62

-----
MULT 12
PRIO_CALJA STANDARD: PRT: 252 AA.
AC P40247;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
GN PRNP.
OS Callitrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callitrix.
OC NCBI_TaxID=9483;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95139066; PubMed=7837269;
RX Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT "Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CRETZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL; U08304; AAC50092.1; -
CC HSSP; P04925; IAG2.
CC InterPro: IPR000817; Prion.
CC Pfam: PF00377; prion.1.
CC PRINTS: PR00341; PRION.
CC SMART; SM00157; PRP.1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT CHAIN 1 22 BY SIMILARITY.
FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 229 MAJOR PRION PROTEIN.
FT PROPEP 230 252 REMOVED IN MATURE FORM (BY SIMILARITY).
FT LIPID 229 229 GPI-ANCHOR (BY SIMILARITY).
FT DISULFID 178 213 BY SIMILARITY.
```

```
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DOMAIN 51 90 5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
FT REPEAT 51 58 1.
FT REPEAT 59 66 2.
FT REPEAT 67 74 3.
FT REPEAT 75 82 4.
FT REPEAT 83 90 5.
SQ SEQUENCE 252 AA; 27639 MW; B2800B60FD5CE664 CRC64;

Query Match 100.0%; Score 55; DB 1; Length 252;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGPHGG 8
Db 55 GWGPHGG 62

-----
RESULT 13
PRIO_CEBAP STANDARD: PRT: 252 AA.
AC P40249;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).
GN PRNP.
OS Cebus apella (Brown-capped capuchin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Cebus.
OC NCBI_TaxID=9515;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95139066; PubMed=7837269;
RX Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;
RT "Prion protein gene variation among primates.";
RL J. Mol. Biol. 245:362-374(1995).
CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND
CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASES KURU,
CC CRETZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL; U08295; AAC50084.1; -
CC HSSP; P04156; IELG.
CC InterPro: IPR000817; Prion.
CC Pfam: PF00377; prion.1.
CC PRINTS: PR00341; PRION.
CC SMART; SM00157; PRP.1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
CC Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT CHAIN 1 22 BY SIMILARITY.
FT SIGNAL 1 22 BY SIMILARITY.
FT CHAIN 23 229 MAJOR PRION PROTEIN.
FT PROPEP 230 252 REMOVED IN MATURE FORM (BY SIMILARITY).
FT LIPID 229 229 GPI-ANCHOR (BY SIMILARITY).
FT DISULFID 178 213 BY SIMILARITY.
```

FT	CARBOHYD	180	180	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	196	196	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	DOMAIN	51	90	5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G
FT				0.
FT	REPEAT	51	58	1.
FT	REPEAT	59	66	2.
FT	REPEAT	67	74	3.
FT	REPEAT	75	82	4.
FT	REPEAT	83	90	5.
SO	SEQUENCE	252 AA:	27579 MM:	A2DFCA0AD26D7821 CRC64:

Query Match	100.0%	Score 55;	DB 1;	Length 252;					
Best Local Similarity	100.0%	Pred. NO. 0.043;							
Matches	8;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0

QY	1	GWGQPHGG	8
Db	55	GWGQPHGG	62

ULT 14	
O_RABIT	
PRIO_RABIT	STANDARD;
	PRT; 252 AA

DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).  
GN Prnp OR Prp.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
NC NCGL\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-New Zealand white;  
RC MDLINE=9718365; PubMed=9031631;  
RX Loftus B., Rogers M.;  
RA "Characterization of a prion protein (PrP) gene from rabbit; a  
RT species with apparent resistance to infection by prions."  
RL Gene 184:215-219(1997).  
-1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE  
-1- HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.  
-1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED  
"RODS".  
-1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
-1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND  
ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASE KURU,  
CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME  
(GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),  
TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.  
-1- SIMILARITY: BELONGS TO THE PRION FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC  
DR EMBL; U28334; AAC48697.1; -.  
DR HSSP; P10279; IDWY.  
DR InterPro; IPR000817; Prion.  
DR Pfam; PF00377; Prion; 1.  
DR PRINTS; PR00341; PRION.  
DR SMART; SM00157; PrP; 1.  
DR PROSITE; PS00291; PRION\_1; 1.  
DR PROSITE; PS00706; PRION\_2; 1.  
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.  
FT SIGNAL 1 28  
FT CHAIN 29 252 MAJOR PRION PROTEIN.  
FT CARBOHYD 180 180 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT

FT	CARBOHYD	196	196	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	DISULFID	178	213	BY SIMILARITY.
FT	DOMAIN	51	92	5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G
FT				0.
FT	REPEAT	51	59	1.
FT	REPEAT	60	67	2.
FT	REPEAT	68	75	3.
FT	REPEAT	76	83	4.
FT	REPEAT	84	92	5.
Q3	SEQUENCE	252 AA;	27432 MF;	ZE177AAFC36B23A54 CRC64;

Query Match	100.0%;	Score 55;	DB 1;	Length 252;
Best Local Similarity	100.0%;	Pred. NO. 0.043;		
Matches . 8;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	GWGQPHGC	8
Db	56	GWGQPHGC	63

RESULT 15	
PRIQ_COLGU	
ID PRIQ_COLGU	STANDARD:
	PRT: 253 AA

DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Major prion protein precursor (PrP) (PrP27-30) (PrP33-35C).  
 GN Prn<sup>P</sup>.  
 OS Colobus guereza (Black-and-white colobus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;  
 OC Colobus.  
 OX NCBI\_taxid=33548;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95139066; PubMed=7837269;  
 RA Schatzl H.M., Dacosta M., Taylor L., Cohen F.E., Prusiner S.B.;  
 RT "Prion protein gene variation among primates."  
 RL J. Mol. Biol. 245:362-374(1995).  
 RN (2)  
 RP SEQUENCE OF 8-253 FROM N.A.  
 RA der Kuyil A.C., Dekker J.T., Goudsmit J.;  
 RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE  
 CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.  
 CC -1- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED  
 CC "RODS".  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
 CC -1- DISEASE: PRP IS FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND  
 CC ANIMALS INFECTED WITH THE DEGENERATIVE NEUROLOGICAL DISEASE KURU,  
 CC CREUTZFELDT-JAKOB DISEASE (CJD), GERSTMAN-STRAUSSLER SYNDROME  
 CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),  
 CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.  
 CC -1- SIMILARITY: BELONGS TO THE PRION FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce@isb-sib.ch>).  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: U08297; AAC50086.1; -;  
 DR EMBL: U75389; AAB50624.1; -;  
 DR HSSP: P04925; IAG2.  
 DR InterPro: IPR000817; Prion.  
 DR Pfam: PF00377; Prion; 1.  
 DR PRINTS: PRO0341; PRION.  
 DR SMART: SMO0157; PRP; 1.  
 DR PROSITE: PS00291; PRION\_1; 1.  
 DR PROSITE: PS00706; PRION\_2; 1.  
 DR

KW Prion: Brain; Glycoprotein; GPI-anchor; Repeat; Signal.  
 FT SIGNAL 1 22 BY SIMILARITY.  
 FT CHAIN 23 230 MAJOR PRION PROTEIN.  
 FT PROPEP 231 253 REMOVED IN MATURE FORM (BY SIMILARITY).  
 FT LIPID 230 230 GPI-ANCHOR (BY SIMILARITY).  
 FT DISULFID 179 214 BY SIMILARITY.  
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT DOMAIN 51 91 5 X 8 AA TANDEM REPEATS OF P-H-G-G-W-G-  
 FT REPEAT 51 59 0.  
 FT REPEAT 60 67 1.  
 FT REPEAT 68 75 2.  
 FT REPEAT 76 83 3.  
 FT REPEAT 84 91 4.  
 SQ SEQUENCE 253 AA; 27626 MW; 14B17477881F5316 CRC64; 5.

Query Match 100.0%; Score 55; DB 1; Length 253;  
 Best Local Similarity 100.0%; Pred. No. 0.043;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHHG 8  
 |||||  
 Db 56 GWGQPHHG 63

Search completed: January 3, 2003, 15:29:18  
 Job time : 7.90909 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:24:35 ; Search time 27.0909 seconds  
(without alignments)  
60.846 Million cell updates/sec

Title: US-09-543-188a-1

Perfect score: 55

Sequence: 1 GWCGRPHG 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

al number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: SPREMBL\_21:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mhcc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	55	100.0	97	4	Q16409 homo sapien
2	55	100.0	105	6	Q97697 muntiacus m
3	55	100.0	117	11	Q9WUJ1 O9WUJ1
4	55	100.0	124	6	Q9TU20 O9TU20
5	55	100.0	141	6	Q97905 tursiops tr
6	55	100.0	143	6	Q9M217 ovis aries
7	55	100.0	181	6	Q97911 budorcas ta
8	55	100.0	185	6	Q97694 cervus nipp
9	55	100.0	195	6	Q97903 addax nasom
10	55	100.0	200	6	Q97693 canis lupus
11	55	100.0	202	6	Q97912 bison bonas
12	55	100.0	202	6	Q97908 capra nubia
13	55	100.0	202	6	Q97696 lama glama
14	55	100.0	204	6	Q97629 odocoileus
15	55	100.0	204	6	Q9TS18 odocoileus
16	55	100.0	204	6	Q9TS17 odocoileus

17	55	100.0	209	6	Q9TV02 camelus dro
18	55	100.0	211	6	Q97787 antilocapra
19	55	100.0	212	6	Q97698 cervus elap
20	55	100.0	213	6	Q9TV04 canis fami1
21	55	100.0	214	6	Q9TV03 canis fami1
22	55	100.0	215	6	Q97904 bos javanic
23	55	100.0	216	6	Q9TV00 bos taurus
24	55	100.0	220	6	Q92825 odocoileus
25	55	100.0	222	6	Q97913 equus quagg
26	55	100.0	223	6	Q97910 hippotragus
27	55	100.0	226	6	Q97907 gazella sub
28	55	100.0	227	6	Q97906 equus cabal
29	55	100.0	227	6	Q97909 tragalaphus
30	55	100.0	227	6	Q97964 equus cabal
31	55	100.0	233	4	P78446 homo sapien
32	55	100.0	235	6	Q97695 giraffa cam
33	55	100.0	240	11	Q8VHV4 microtus ag
34	55	100.0	243	11	P97895 mesocricetu
35	55	100.0	245	4	Q15216 homo sapien
36	55	100.0	245	6	Q9M2U7 mus musculu
37	55	100.0	246	4	O60489 homo sapien
38	55	100.0	248	11	Q8VHV5 clethrionom
39	55	100.0	253	4	Q9UP19 homo sapien
40	55	100.0	253	4	Q96E70 homo sapien
41	55	100.0	253	4	Q8TR60 homo sapien
42	55	100.0	253	11	Q9Z075 meriones un
43	55	100.0	254	6	Q9TSE8 oryctolagus
44	55	100.0	254	11	Q9Z0T4 sigmodon fu
45	55	100.0	254	11	Q9QYT9 mus musculu

#### ALIGNMENTS

RESULT 1  
ID 016409 PRELIMINARY; PRT; 97 AA.  
AC Q16409; Q16407.  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE PPP amyloid (Prion protein) (Fragment).  
GN PRNP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP MEDLINE=92140671; PubMed=1736177;  
RA Brown P., Goldfarb L.G., McCombie W.R., Nieto A., Squillacote D.,  
RT Sheremata W., Little B.W., Godec M.S., Gibbs C.J., Jr., Gajdusek D.C.;  
RT "Atypical Creutzfeldt-Jakob disease in an American family with an  
RT insert mutation in the PRNP amyloid precursor gene.";  
RL Neurology 42:422-427(1992).  
RN [2]  
RP SEQUENCE OF 1-89 FROM N.A.  
RX MEDLINE=96057178; PubMed=7572084;  
RA Oda T., Kitamoto T., Tateishi J., Mitsuhashi T., Iwabuchi K., Haga C.,  
RT Oguni E., Kato Y., Tomimaga I., Yanai K.;  
RT "Prion disease with 144 base pair insertion in a Japanese family  
RT line.";  
RL Acta Neuropathol. 90:80-86(1995).  
RN [3]  
RP SEQUENCE OF 1-81 FROM N.A.  
RX MEDLINE=96390485; PubMed=8797471;  
RA Cochran E.J., Bennett D.A., Cervenkova L., Kenney K., Bernard B.,  
RT "Familial Creutzfeldt-Jakob disease with a five-repeat octapeptide  
RT insert mutation.";  
RL Neurology 47:727-733(1996).  
DR EMBL: S80539; AAB21334.1; -;  
EMBL: S79978; AAB35416.1; -;

DR EMBL: S83341; AAB50777.1; -  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; Prion; 2.  
DR PRINTS: PR00341; PRION.  
KW Prion.  
FT NON\_TER 1 1  
SQ SEQUENCE 97 AA; 9388 MW; D2A9F370DD10511D CRC64;

Query Match 100.0%; Score 55; DB 4; Length 97;  
Best Local Similarity 100.0%; Pred. No. 0.079;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWCOPHCG 8  
Db 46 GWCOPHCG 53

RESULT 2  
097697 PRELIMINARY; PRT; 105 AA.

097697;  
01-MAY-1999 (TREMBLrel. 10, Created)  
01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Prion protein (Fragment).  
GN PRP.  
OS Muntiacus muntjak (Muntjak).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae;  
OC Cervidae; Muntiacinae; Muntiacus.  
OX NCBI\_TaxID=9888;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=97317556; PubMed=9174569;  
RA Scharzl H.M., Wopfinger F., Gilch S., von Brunn A., Jager G.;  
RT "Is codon 129 of prion protein polymorphic in human beings but not in  
animal?"  
RL Lancet 349:1603-1604(1997).  
RN [2]

RP SEQUENCE FROM N.A.  
RX MEDLINE=99303687; PubMed=10373359;  
RA Wopfinger F., Weldenhofer G., Schneider R., von Brunn A., Gilch S.,  
RA Schwarz T.F., Werner T., Scharzl H.M.;  
RT "Analysis of 27 mammalian and 9 avian prps reveals high conservation  
of flexible regions of the prion protein.";  
RL J. Mol. Biol. 289:1163-1178(1999).  
DR EMBL:AF113944; AAD13292.1; -  
DR InterPro: IPR002395; Kininogen.

DR Pfam: PF00377; Prion; 1.  
DR PRINTS: PR00341; KININOGEN.  
DR PRINTS: PR00341; PRION.  
DR SMART: SM00157; PRP; 1.  
FT NON\_TER 1 1  
FT NON\_TER 105 105  
SQ SEQUENCE 105 AA; 10539 MW; 8F88FE2AE10C693 CRC64;

Query Match 100.0%; Score 55; DB 6; Length 105;  
Best Local Similarity 100.0%; Pred. No. 0.086;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWCOPHCG 8  
Db 51 GWCOPHCG 58

RESULT 3  
09WUJ1 PRELIMINARY; PRT; 117 AA.

AC 09WUJ1;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Prion protein (Fragment).  
OS Cavia porcellus (Guinea pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae; Cavia.  
OX NCBI\_TaxID=10141;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=STRAIN 13/N; TISSUE=OCULAR LENS;  
RA Frederikse P.F., Zigler J.S. Jr., Carper D.A.;  
RT "Expression of Prion Protein in Human Lenses."  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF139166; AAD24491.1; -  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; Prion; 1.  
DR PRINTS: PR00341; PRION.  
DR SMART: SM00157; PRP; 1.  
DR PROSITE: PS00291; PRION\_1; 1.  
FT NON\_TER 1 1  
FT NON\_TER 117 117  
SQ SEQUENCE 117 AA; 11818 MW; 863A02472404EA05 CRC64;

Query Match 100.0%; Score 55; DB 11; Length 117;  
Best Local Similarity 100.0%; Pred. No. 0.096;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWCOPHCG 8  
Db 54 GWCOPHCG 61

RESULT 4  
09TU20 PRELIMINARY; PRT; 124 AA.

AC 09TU20;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Prion protein (Fragment).  
GN PRP.

OS Varecia variegata variegata.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Strepsirhini; Lemnidae; Varecia.  
OX NCBI\_TaxID=87289;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gilch S., Scharzl H.M.;  
RT "Unusual prion protein octarepeat structure of the highly BSE-  
susceptible lemur monkey.";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF177293; AAD54335.1; -  
DR HSSP: P04925; IAG2.  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; Prion; 1.  
DR PRINTS: PR00341; PRION.  
DR SMART: SM00157; PRP; 1.  
DR PROSITE: PS00291; PRION\_1; 1.  
FT NON\_TER 1 1  
FT NON\_TER 124 124  
SQ SEQUENCE 124 AA; 13436 MW; CC2C8A5A855A7C94 CRC64;

Query Match 100.0%; Score 55; DB 6; Length 124;  
Best Local Similarity 100.0%; Pred. No. 0.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWCOPHCG 8  
Db 5 GWCOPHCG 12

RESULT 5  
097905 PRELIMINARY; PRT; 141 AA.  
AC 097905;



```
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Prion protein (Fragment).
GN PRP.
OS Tursiops truncatus (Atlantic bottle-nosed dolphin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
OC Tursiops.
OC NCBI_TaxID=9739;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=PBL;
RX MEDLINE=99303687; PubMed=10373359;
RA Wopfner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,
Schwarz T.F., Werner T., Scharzl H.M.;
RT "Analysis of 27 mammalian and 9 avian PRPs reveals high conservation
of flexible regions of the prion protein."
RL J. Mol. Biol. 289:1163-1178(1999).
OC EMBL: AF117311; AAD19982.1; -.
HSSP: P10279; IDMY.
DR InterPro: IPR000817; Prion.
DR Pfam: PF00377; Prion.
DR PRINTS: PR00341; PRION.
DR SMART: SM00157; PRP. 1.
DR PROSITE: PS00291; PRION_1; 1.
FT NON_TER 1 1
FT NON_TER 141 141
SQ SEQUENCE 141 AA; 15066 MW; 79BE306E2AAA187C CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 6; Length 141;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GWGQPHGC 8
Db 36 GWGQPHGC 43

RESULT 6
O9M217 PRELIMINARY; PRT; 143 AA.
AC O9M217;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Prion protein (Fragment).
OC Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OC NCBI_TaxID=9940;
RN (1)
RP SEQUENCE FROM N.A.
RA Herrmann L.M., Davis W.B., Li H., Wardrop J., Sy M.-S., Gambetti P.,
Knowles D.P.;
RT "Detection of PrPSc in peripheral blood mononuclear cells of scrapie
susceptible sheep."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF267507; AAF91403.1; -.
HSSP: P10279; IDMY.
DR InterPro: IPR000817; Prion.
DR Pfam: PF00377; Prion.
DR PRINTS: PR00341; PRION.
DR SMART: SM00157; PRP. 1.
DR PROSITE: PS00291; PRION_1; 1.
FT NON_TER 143 143
FT NON_TER 143 143
SQ SEQUENCE 143 AA; 14428 MW; 4E2D296C6C8022E2 CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 6; Length 143;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 GWGQPHGC 8
Db 59 GWGQPHGC 66

RESULT 7
O97911 PRELIMINARY; PRT; 181 AA.
AC O97911;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Prion protein (Fragment).
GN PRP.
OS Budorcas taxicolor (takin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Budorcas.
OC NCBI_TaxID=37181;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=PBL;
RX MEDLINE=99303687; PubMed=10373359;
RA Wopfner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,
Schwarz T.F., Werner T., Scharzl H.M.;
RT "Analysis of 27 mammalian and 9 avian PRPs reveals high conservation
of flexible regions of the prion protein."
RL J. Mol. Biol. 289:1163-1178(1999).
OC EMBL: AF117326; AAD19997.1; -.
HSSP: P10279; IDMY.
DR InterPro: IPR002395; Kininogen.
DR InterPro: IPR000817; Prion.
DR Pfam: PF00377; Prion.
DR PRINTS: PR00341; KININOGEN.
DR PRINTS: PR00341; PRION.
DR SMART: SM00157; PRP. 1.
DR PROSITE: PS00291; PRION_1; 1.
FT NON_TER 1 1
FT NON_TER 181 181
SQ SEQUENCE 181 AA; 19253 MW; A9001D086442E92A CRC64;

Query Match
Best Local Similarity 100.0%; Score 55; DB 6; Length 181;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GWGQPHGC 8
Db 32 GWGQPHGC 39

RESULT 8
O97694 PRELIMINARY; PRT; 185 AA.
AC O97694;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Prion protein (Fragment).
GN PRP.
OS Ceruus nippon dybowskii.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC Cervidae; Cervinae; Cervus.
OC NCBI_TaxID=88066;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=97317556; PubMed=9174569;
RA Scharzl H.M., Wopfner F., Gilch S., von Brunn A., Jager G.;
RT "Is codon 129 of prion protein polymorphic in human beings but not in
animals?"
RL Lancet 349:1603-1604(1997).
RN (2)
RP SEQUENCE FROM N.A.
```

RX MEDLINE=99303687; PubMed=10373359;  
RA Wopfner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,  
Schwarz T.F., Werner T., Schatzl H.M.;  
RT "Analysis of 27 mammalian and 9 avian PrPs reveals high conservation  
of flexible regions of the prion protein.";  
RL J. Mol. Biol. 289:1163-1178(1999).  
DR EMBL: AF113941; AAD13289.1; -  
DR HSSP: P10279; 1DWY.  
DR InterPro: IPR002395; Kininogen.  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; Prion; 1.  
DR PRINTS: PR00344; KININOGEN.  
DR PRINTS: PR00341; PRION.  
DR SMART: SM00157; PRP; 1.  
DR PROSITE: PS00291; PRION\_1; 1.  
FT NON\_TER 1 185  
FT SEQUENCE 185 AA; 19870 MW; BB87C7658BC66E79 CRC64;

Query Match 100.0%; Score 55; DB 6; Length 185;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGPHHG 8  
DB 35 GWGPHHG 42

RESULT 9

ID 097903 PRELIMINARY; PRT; 195 AA.  
AC 097903;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Prion protein (Fragment).  
GN PRP.  
OS Addax nasomaculatus.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Hippotraginae; Addax.  
OX NCBI\_TaxID=35915;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PBL.  
RX MEDLINE=99303687; PubMed=10373359;

RA Wopfner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,  
Schwarz T.F., Werner T., Schatzl H.M.;  
RT "Analysis of 27 mammalian and 9 avian PrPs reveals high conservation  
of flexible regions of the prion protein.";  
RL J. Mol. Biol. 289:1163-1178(1999).  
DR EMBL: AF117309; AAD19980.1; -  
DR HSSP: P10279; 1DWY.  
DR InterPro: IPR002395; Kininogen.  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; Prion; 1.  
DR PRINTS: PR00344; KININOGEN.  
DR PRINTS: PR00341; PRION.  
DR SMART: SM00157; PRP; 1.  
DR PROSITE: PS00291; PRION\_1; 1.  
DR PROSITE: PS00706; PRION\_2; 1.  
FT NON\_TER 1 195  
FT SEQUENCE 195 AA; 21331 MW; 6A9BA6A7E1AFCA9 CRC64;

Query Match 100.0%; Score 55; DB 6; Length 195;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGPHHG 8  
DB 23 GWGPHHG 30

RESULT 10  
ID 097693 PRELIMINARY; PRT; 195 AA.  
AC 097693;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Prion protein (Fragment).  
GN PRP.  
OS Canis lupus (Gray wolf).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9612;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99303687; PubMed=10373359;  
RA Wopfner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,  
Schwarz T.F., Werner T., Schatzl H.M.;  
RT "Analysis of 27 mammalian and 9 avian PrPs reveals high conservation  
of flexible regions of the prion protein.";  
RL J. Mol. Biol. 289:1163-1178(1999).  
DR EMBL: AF113939; AAD12063.1; -  
DR HSSP: P04925; 1AG2.  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; Prion; 1.  
DR PRINTS: PR00341; PRION.  
DR SMART: SM00157; PRP; 1.  
DR PROSITE: PS00291; PRION\_1; 1.  
DR PROSITE: PS00706; PRION\_2; 1.  
FT NON\_TER 1 195  
FT SEQUENCE 195 AA; 21097 MW; 9D18E4E9AA5D031 CRC64;

Query Match 100.0%; Score 55; DB 6; Length 195;  
Best Local Similarity 100.0%; Pred. No. 0.16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGPHHG 8  
DB 18 GWGPHHG 25

RESULT 11

ID 097912 PRELIMINARY; PRT; 200 AA.  
AC 097912;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Prion protein (Fragment).  
GN PRP.  
OS Bison bonasus (European bison).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bison.  
OX NCBI\_TaxID=9902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PBL.  
RX MEDLINE=99303687; PubMed=10373359;

RA Wopfner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,  
Schwarz T.F., Werner T., Schatzl H.M.;  
RT "Analysis of 27 mammalian and 9 avian PrPs reveals high conservation  
of flexible regions of the prion protein.";  
RL J. Mol. Biol. 289:1163-1178(1999).  
DR EMBL: AF117328; AAD19999.1; -  
DR HSSP: P10279; 1DWY.  
DR InterPro: IPR002395; Kininogen.  
DR InterPro: IPR001610; PAC.  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; Prion; 1.  
DR PRINTS: PR00344; KININOGEN.

DR PRINTS; PR00341; PRION.  
DR SMART; SM00086; PAC; 1.  
DR SMART; SM00157; PRP; 1.  
DR PROSITE; PS00291; PRION\_1; 1.  
DR PROSITE; PS00706; PRION\_2; 1.  
FT NON\_TER 1 1  
FT NON\_TER 200 200  
SQ SEQUENCE 200 AA; 21674 MW; 1F270CDF4BE5271B CRC64;

Query Match  
Best Local Similarity 100.0%; Score 55; DB 6; Length 200;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHG 8  
DB 16 GWGQPHG 23

RESULT 12

OY 097908 PRELIMINARY; PRT: 202 AA.  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Prion protein (Fragment).  
GN PRP.

OS Capra nubiana (Nubian ibex).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Capra.  
OX NCBI\_TaxID=72543;  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE-PBL;  
RA MEDLINE-99303687; PubMed-10373359;  
RA Wopner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,  
RA Schwarz T.F., Werner T., Scharz H.M.;  
RT "Analysis of 27 mammalian and 9 avian PRPs reveals high conservation  
of flexible regions of the prion protein.";  
RL J. Mol. Biol. 289:1163-1178(1999).

DR EMBL; AF117319; AAD19990.1; -.  
DR HSSP; P10279; IDWY.  
DR InterPro; IPR002395; Kininogen.  
DR InterPro; IPR000817; Prion.  
DR Pfam; PF00377; prion; 1.  
DR PRINTS; PR00341; KININOGEN.  
DR PRINTS; PR00341; PRION.  
DR SMART; SM00157; PRP; 1.  
DR PROSITE; PS00291; PRION\_1; 1.  
DR PROSITE; PS00706; PRION\_2; 1.  
FT NON\_TER 1 1  
FT NON\_TER 202 202  
SQ SEQUENCE 202 AA; 21949 MW; DB0634A43B4DB77F CRC64;

Query Match  
Best Local Similarity 100.0%; Score 55; DB 6; Length 202;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHG 8  
DB 31 GWGQPHG 38

RESULT 13

OY 097696 PRELIMINARY; PRT: 202 AA.  
AC 097696;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Prion protein (Fragment).  
GN PRP.

OS Lama glama (Llama).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.  
OX NCBI\_TaxID=9844;  
RN [1]  
RP SEQUENCE FROM N.A.

RA MEDLINE-99303687; PubMed-10373359;  
RA Wopner F., Weidenhofer G., Schneider R., von Brunn A., Gilch S.,  
RA Schwarz T.F., Werner T., Scharz H.M.;  
RT "Analysis of 27 mammalian and 9 avian PRPs reveals high conservation  
of flexible regions of the prion protein.";  
RL J. Mol. Biol. 289:1163-1178(1999).

DR EMBL; AF113943; AAD13291.1; -.  
DR HSSP; P10279; IDWY.  
DR InterPro; IPR000817; Prion.  
DR Pfam; PF00377; prion; 1.  
DR PRINTS; PR00341; PRION.  
DR SMART; SM00157; PRP; 1.  
DR PROSITE; PS00291; PRION\_1; 1.  
DR PROSITE; PS00706; PRION\_2; 1.  
FT NON\_TER 1 1  
FT NON\_TER 202 202  
SQ SEQUENCE 202 AA; 21860 MW; FC45232DB773F354 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 55; DB 6; Length 202;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHG 8  
DB 11 GWGQPHG 18

RESULT 14

OY 097629 PRELIMINARY; PRT: 204 AA.  
AC 097629;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Prion protein (Fragment).  
GN PRP.

OS Odocoileus virginianus (White-tailed deer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;  
OC Cervidae; Odocoileinae; Odocoileus.  
OX NCBI\_TaxID=9874;  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE-BRAIN;  
RA O'Rourke K.I., Miller M.W., Wild M.A., Williams E.S.;  
RT "Prp alleles in free ranging and captive white tailed deer (Odocoileus  
virginianus).";  
RT Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.  
RL EMBL; AF091558; AAC69626.1; -.

DR HSSP; P10279; IDWY.  
DR InterPro; IPR002395; Kininogen.  
DR InterPro; IPR000817; Prion.  
DR Pfam; PF00377; prion; 1.  
DR PRINTS; PR00341; KININOGEN.  
DR PRINTS; PR00341; PRION.  
DR SMART; SM00157; PRP; 1.  
DR PROSITE; PS00291; PRION\_1; 1.  
DR PROSITE; PS00706; PRION\_2; 1.  
FT NON\_TER 1 1  
FT NON\_TER 204 204  
SQ SEQUENCE 204 AA; 22154 MW; CABA68F2B49C81E CRC64;

Query Match  
Best Local Similarity 100.0%; Score 55; DB 6; Length 204;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHG 8

Db 36 GWGQPHGG 43

## RESULT 15

09TS18  
ID 09TS18 PRELIMINARY: PRT: 204 AA.  
AC 09TS18;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Prion protein (Fragment).  
GN PRP.  
OS Odocoileus virginianus (White-tailed deer).  
OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;  
OC Cervidae; Odocoileinae; Odocoileus.  
OX / NCBI\_TaxID=9874;  
RN (1)  
SEQUENCE FROM N.A.

TISSUE-BRAIN;  
O'Rourke K.I., Miller M.W., Wild M.A., Williams E.S.;  
"Prp alleles in free ranging and captive white tailed deer (Odocoileus  
virginianus).";  
Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF091559; AAC69627.1; -  
DR HSSP: P10279; IDWY.  
DR InterPro: IPR000817; Prion.  
DR Pfam: PF00377; Prion; 1.  
DR PRINTS: PR00341; PRION.  
DR SMART: SM00157; PRP; 1.  
DR PROSITE: PS00291; PRION\_1; 1.  
DR PROSITE: PS00706; PRION\_2; 1.  
FT NON\_TER 1 1  
FT NON\_TER 204 204  
SQ SEQUENCE 204 AA; 22181 MW; CA962B93FA84D4D3 CRC64;

Query Match 100.0%; Score 55; DB 6; Length 204;  
Best Local Similarity 100.0%; Pred. NO. 0.17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8  
DB 36 GWGQPHGG 43

Search completed: January 3, 2003, 15:31:54  
Job time : 28.0909 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:16 : Search time 12 seconds  
(without alignments)  
19.615 Million cell updates/sec

Title: US-09-543-188A-1  
Perfect score: 55  
Sequence: 1 GWGQPHGC 8

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues  
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :  
1: Issued\_Patents\_AA:\*  
2: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	18	1	US-08-244-701B-24
2	55	100.0	18	1	US-08-244-701B-27
3	55	100.0	18	1	US-08-244-701B-30
4	55	100.0	18	4	US-09-076-721-24
5	55	100.0	18	4	US-09-076-721-27
6	55	100.0	18	4	US-09-076-721-30
7	55	100.0	26	1	US-08-244-701B-46
8	55	100.0	26	4	US-09-076-721-46
9	55	100.0	208	4	US-09-128-450-18
10	55	100.0	208	4	US-09-823-494-18
11	55	100.0	253	1	US-08-242-188-2
12	55	100.0	253	1	US-08-509-261A-2
13	55	100.0	253	1	US-08-660-626-8
14	55	100.0	253	1	US-08-692-892-2
15	55	100.0	253	2	US-08-713-939A-2
16	55	100.0	253	2	US-08-868-162A-22
17	55	100.0	253	4	US-09-031-168-8
18	55	100.0	253	4	US-09-128-450-20
19	55	100.0	253	4	US-09-036-579-2
20	55	100.0	253	4	US-09-823-494-20
21	55	100.0	253	4	US-09-550-374-2
22	55	100.0	254	1	US-08-242-188-1
23	55	100.0	254	1	US-08-509-261A-1
24	55	100.0	254	1	US-08-660-626-7
25	55	100.0	254	1	US-08-692-892-1
26	55	100.0	254	2	US-08-713-939A-1
27	55	100.0	254	2	US-08-868-162A-21

28	55	100.0	254	4	US-09-031-168-7	Sequence 7, Appli
29	55	100.0	254	4	US-09-128-450-19	Sequence 16, Appl
30	55	100.0	254	4	US-09-128-450-26	Sequence 26, Appl
31	55	100.0	254	4	US-09-128-450-28	Sequence 2, Appl
32	55	100.0	254	4	US-09-036-579-1	Sequence 1, Appli
33	55	100.0	254	4	US-09-823-494-19	Sequence 19, Appl
34	55	100.0	254	4	US-09-823-494-26	Sequence 26, Appl
35	55	100.0	254	4	US-09-823-494-28	Sequence 28, Appl
36	55	100.0	254	4	US-09-550-374-1	Sequence 1, Appli
37	55	100.0	255	1	US-08-242-188-4	Sequence 4, Appli
38	55	100.0	255	1	US-08-509-261A-4	Sequence 4, Appli
39	55	100.0	255	1	US-08-660-626-10	Sequence 10, Appl
40	55	100.0	255	1	US-08-692-892-4	Sequence 4, Appli
41	55	100.0	255	2	US-08-713-939A-4	Sequence 4, Appli
42	55	100.0	255	2	US-08-868-162A-24	Sequence 24, Appl
43	55	100.0	255	4	US-09-031-168-10	Sequence 10, Appl
44	55	100.0	255	4	US-09-036-579-4	Sequence 4, Appli
45	55	100.0	255	4	US-09-550-374-4	Sequence 4, Appli

## ALIGNMENTS

RESULT 1  
US-08-244-701B-24  
Sequence 24, Application US/08244701B  
Patent No. 3773572  
GENERAL INFORMATION:  
APPLICANT: Fishleigh, Robert V.  
APPLICANT: Robson, Barry  
APPLICANT: Mee, Roger P.  
TITLE OF INVENTION: Fragments of prion proteins  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESS: Pennine & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/244,701B  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Fanucci, Allan A.  
REGISTRATION NUMBER: 30,256  
REFERENCE/DOCKET NUMBER: 8080-007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /label= X  
OTHER INFORMATION: /note= "X may be absent or present independently  
OTHER INFORMATION: of Y and denotes one or more amino acid(s)"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 18

OTHER INFORMATION: /label-y  
OTHER INFORMATION: /note="y may be absent or present independently  
OTHER INFORMATION: of x and denotes one or more amino acid(s)"  
US-08-244-701B-24

Query Match 100.0%; Score 55; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGPHHG 8  
Db 4 GWGPHHG 11

## RESULT 2

US-08-244-701B-27  
Sequence 27, Application US/08244701B  
Patent No. 5773572

## GENERAL INFORMATION:

APPLICANT: Fishleigh, Robert V.

APPLICANT: Robson, Barry

APPLICANT: Mee, Roger P.

TITLE OF INVENTION: Fragments of Prion Proteins

NUMBER OF SEQUENCES: 67

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/244,701B

FILING DATE: 02-JUN-1994

CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:

NAME: Fanucci, Allan A.

REGISTRATION NUMBER: 30,256

REFERENCE/DOCKET NUMBER: 8080-007

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1

OTHER INFORMATION: /label-x

OTHER INFORMATION: /note="x may be absent or present independently

OTHER INFORMATION: of y and denotes one or more amino acid(s)"

FEATURE:

NAME/KEY: Modified-site

LOCATION: 18

OTHER INFORMATION: /label-y

OTHER INFORMATION: /note="y may be absent or present independently

OTHER INFORMATION: of x and denotes one or more amino acid(s)"

US-08-244-701B-27

OY 1 GWGPHHG 8  
Db 4 GWGPHHG 11

Query Match 100.0%; Score 55; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGPHHG 8  
Db 4 GWGPHHG 11

## RESULT 3

US-08-244-701B-30  
Sequence 30, Application US/08244701B  
Patent No. 5773572

## GENERAL INFORMATION:

APPLICANT: Fishleigh, Robert V.

APPLICANT: Robson, Barry

APPLICANT: Mee, Roger P.

TITLE OF INVENTION: Fragments of Prion Proteins

NUMBER OF SEQUENCES: 67

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/244,701B

FILING DATE: 02-JUN-1994

CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:

NAME: Fanucci, Allan A.

REGISTRATION NUMBER: 30,256

REFERENCE/DOCKET NUMBER: 8080-007

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1

OTHER INFORMATION: /label-x

OTHER INFORMATION: /note="x is absent or present independently of y

OTHER INFORMATION: and denotes one or more amino acid(s)"

FEATURE:

NAME/KEY: Modified-site

LOCATION: 18

OTHER INFORMATION: /label-y

OTHER INFORMATION: /note="y may be absent or present independently

OTHER INFORMATION: of x and denotes one or more amino acid(s)"

US-08-244-701B-30

OY 1 GWGPHHG 8  
Db 4 GWGPHHG 11

Query Match 100.0%; Score 55; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGPHHG 8  
Db 4 GWGPHHG 11

## RESULT 4

US-09-076-721-24  
Sequence 24, Application US/09076721  
Patent No. 6379905

## GENERAL INFORMATION:

APPLICANT: Fishleigh, Robert V.

APPLICANT: Robson, Barry

APPLICANT: Mee, Roger P.

TITLE OF INVENTION: Fragments of Prion Proteins

NUMBER OF SEQUENCES: 67

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/244,701B

FILING DATE: 02-JUN-1994

CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:

NAME: Fanucci, Allan A.

REGISTRATION NUMBER: 30,256

REFERENCE/DOCKET NUMBER: 8080-007

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-8864/9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1

OTHER INFORMATION: /label-x

OTHER INFORMATION: /note="x is absent or present independently of y

OTHER INFORMATION: and denotes one or more amino acid(s)"

FEATURE:

NAME/KEY: Modified-site

LOCATION: 18

OTHER INFORMATION: /label-y

OTHER INFORMATION: /note="y may be absent or present independently

OTHER INFORMATION: of x and denotes one or more amino acid(s)"

US-08-244-701B-30

```

; GENERAL INFORMATION:
; APPLICANT: Fishleigh, Robert V.
; APPLICANT: Robson, Barry
; APPLICANT: Mee, Roger P.
; TITLE OF INVENTION: Fragments of Prion Proteins
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,721
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/244,701
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 8080-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /label=X
; OTHER INFORMATION: /note="X may be absent or present independently
; OTHER INFORMATION: of Y and denotes one or more amino acid(s)"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 18
; OTHER INFORMATION: /label=Y
; OTHER INFORMATION: /note="Y may be absent or present independently
; OTHER INFORMATION: of X and denotes one or more amino acid(s)"
; US-09-076-721-24
;
; Query Match 100.0%; Score 55; DB 4; Length 18;
; Best Local Similarity 100.0%; Pred. No. 0.017;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 GWGQPHGC 8
; Db 4 GWGQPHGC 11
;
; RESULT 5
; US-09-076-721-27
; Sequence 27, Application US/09076721
; Patent No. 6379905
; GENERAL INFORMATION:
; APPLICANT: Fishleigh, Robert V.
; APPLICANT: Robson, Barry
; APPLICANT: Mee, Roger P.
; TITLE OF INVENTION: Fragments of Prion Proteins
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/076,721
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/244,701
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 8080-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /label=X
; OTHER INFORMATION: /note="X may be absent or present independently
; OTHER INFORMATION: of Y and denotes one or more amino acid(s)"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 18
; OTHER INFORMATION: /label=Y
; OTHER INFORMATION: /note="Y may be absent or present independently
; OTHER INFORMATION: of X and denotes one or more amino acid(s)"
; US-09-076-721-27
;
; Query Match 100.0%; Score 55; DB 4; Length 18;
; Best Local Similarity 100.0%; Pred. No. 0.017;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 GWGQPHGC 8
; Db 4 GWGQPHGC 11
;
; RESULT 6
; US-09-076-721-30
; Sequence 30, Application US/09076721
; Patent No. 6379905
; GENERAL INFORMATION:
; APPLICANT: Fishleigh, Robert V.
; APPLICANT: Robson, Barry
; APPLICANT: Mee, Roger P.
; TITLE OF INVENTION: Fragments of Prion Proteins
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
```

ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/076,721  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/244,701  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Panucci, Allan A.  
REGISTRATION NUMBER: 30,256  
REFERENCE/DOCKET NUMBER: 8080-007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNTE  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1  
OTHER INFORMATION: /label=X  
OTHER INFORMATION: /note="X is absent or present independently of Y  
OTHER INFORMATION: and denotes one or more amino acid(s)"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 18  
OTHER INFORMATION: /label=Y  
OTHER INFORMATION: /note="Y may be absent or present independently  
OTHER INFORMATION: of X and denotes one or more amino acid(s)"  
US-09-076-721-30  
Query Match 100.0%; Score 55; DB 4; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GWCOPHGG 8  
11111111  
4 GWCOPHGG 11  
RESULT 7  
US-08-244-701B-46  
Sequence 46, Application US/08244701B  
Patent No. 5773572  
GENERAL INFORMATION:  
APPLICANT: Fishleigh, Robert V.  
APPLICANT: Robson, Barry  
APPLICANT: Mee, Roger P.  
TITLE OF INVENTION: Fragments of Prion Proteins  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/244,701B  
FILING DATE: 02-JUN-1994  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Panucci, Allan A.  
REGISTRATION NUMBER: 30,256  
REFERENCE/DOCKET NUMBER: 8080-007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNTE  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-244-701B-46  
Query Match 100.0%; Score 55; DB 1; Length 26;  
Best Local Similarity 100.0%; Pred. No. 0.024;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GWCOPHGG 8  
11111111  
Db 7 GWCOPHGG 14  
RESULT 8  
US-09-076-721-46  
Sequence 46, Application US/09076721  
Patent No. 6379905  
GENERAL INFORMATION:  
APPLICANT: Fishleigh, Robert V.  
APPLICANT: Robson, Barry  
APPLICANT: Mee, Roger P.  
TITLE OF INVENTION: Fragments of Prion Proteins  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/076,721  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/244,701  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Panucci, Allan A.  
REGISTRATION NUMBER: 30,256  
REFERENCE/DOCKET NUMBER: 8080-007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNTE  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear



MOLECULE TYPE: peptide  
US-09-076-721-46

Query Match 100.0%; Score 55; DB 4; Length 26;  
Best Local Similarity 100.0%; Pred. No. 0.024;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8  
|||||||  
DB 7 GWGQPHGG 14

RESULT 9  
US-09-128-450-18

; Sequence 18, Application US/09128450

; Patent No. 6211149

; GENERAL INFORMATION:

; APPLICANT: Chesebri, Bruce W

; APPLICANT: Caughey, Byron W

; APPLICANT: Chabry, Joelle

; APPLICANT: Priola, Susette

; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion

; FILE REFERENCE: 50121

; CURRENT APPLICATION NUMBER: US/09/128,450

; CURRENT FILING DATE: 1998-08-03

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 18

; LENGTH: 208

; TYPE: PRT

; ORGANISM: Hamster sp.

; FEATURE:

; NAME/KEY: NON\_TER

; LOCATION: (1)..(2)

; US-09-128-450-18

Query Match 100.0%; Score 55; DB 4; Length 208;  
Best Local Similarity 100.0%; Pred. No. 0.18;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8  
|||||||  
DB 34 GWGQPHGG 41

RESULT 10

US-09-823-494-18

; Sequence 18, Application US/09823494

; Patent No. 6355610

; GENERAL INFORMATION:

; APPLICANT: Chesebri, Bruce W

; APPLICANT: Caughey, Byron W

; APPLICANT: Chabry, Joelle

; APPLICANT: Priola, Susette

; TITLE OF INVENTION: Inhibitors of Formation of Protease Resistant Prion

; FILE REFERENCE: 50121

; CURRENT APPLICATION NUMBER: US/09/823,494

; CURRENT FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: 09/128,450

; PRIOR FILING DATE: 1998-08-03

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 18

; LENGTH: 208

; TYPE: PRT

; ORGANISM: Hamster sp.

; FEATURE:

; NAME/KEY: NON\_TER

; LOCATION: (1)..(2)

; US-09-823-494-18

Query Match 100.0%; Score 55; DB 4; Length 208;

Best Local Similarity 100.0%; Pred. No. 0.18;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8  
|||||||  
DB 34 GWGQPHGG 41

RESULT 11

US-08-242-188-2

; Sequence 2, Application US/08242188

; Patent No. 5565186

; GENERAL INFORMATION:

; APPLICANT: Prusiner, Stanley B.

; APPLICANT: Scott, Michael R.

; APPLICANT: Telling, Glenn

; TITLE OF INVENTION: METHOD OF DETECTING PRIONS IN A SAMPLE

; TITLE OF INVENTION: AND TRANSGENIC ANIMAL USED FOR SAME

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Karl Bosicevic

; STREET: 2200 Sand Hill Road

; CITY: Menlo Park

; STATE: CA

; COUNTRY: USA

; ZIP: 94025

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/242,188

; FILING DATE: 13-MAY-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Bosicevic, Karl

; REGISTRATION NUMBER: 28,807

; REFERENCE/DOCKET NUMBER: 06510/014001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 854-5277

; TELEFAX: (415) 854-0875

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 253 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; ORIGINAL SOURCE:

; ORGANISM: HUMAN PRION PROTEIN, HuPrP

; US-08-242-188-2

Query Match 100.0%; Score 55; DB 1; Length 253;

Best Local Similarity 100.0%; Pred. No. 0.22;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8  
|||||||  
DB 56 GWGQPHGG 63

RESULT 12

US-08-509-261A-2

; Sequence 2, Application US/08509261A

; Patent No. 5763244

; GENERAL INFORMATION:

; APPLICANT: Prusiner, Stanley B.

; APPLICANT: Scott, Michael R.

; APPLICANT: Telling, Glenn

; TITLE OF INVENTION: Method of Detecting Prions

; TITLE OF INVENTION: in a Sample and Transgenic Animal Used fore

NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bozicevic & Reed, LLP  
STREET: 285 Hamilton Avenue, Suite 200  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/509,261A  
FILING DATE: 31-JUL-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bozicevic, Karl  
REGISTRATION NUMBER: 28, 807  
REFERENCE/DOCKET NUMBER: 6510-030001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-327-3400  
TELEFAX: 650 327-3231  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-509-261A-2

Query Match 100.0%; Score 55; DB 1; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8  
|||||  
DB 56 GWGQPHGG 63

RESULT 13  
US-08-660-626-8  
Sequence 8, Application US/08660626  
Patent No. 5789655  
GENERAL INFORMATION:  
APPLICANT: Stanley B. Prusiner  
APPLICANT: Glenn C. Telling  
APPLICANT: Fred E. Cohen  
APPLICANT: Michael R. Scott  
TITLE OF INVENTION: TRANSGENIC ANIMALS EXPRESSING  
TITLE OF INVENTION: EPILOPE-TAGGED PROTEINS  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 2200 Sand Hill Road, Suite 100  
CITY: Menlo Park  
STATE: California  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Ascliti  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/660,626  
FILING DATE:  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Valela Gregg  
REGISTRATION NUMBER: 35,127  
REFERENCE/DOCKET NUMBER: 07532/003001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 322-5070  
TELEFAX: (415) 854-0875  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: HUMAN PRION PROTEIN, HuPrP  
US-08-660-626-8

Query Match 100.0%; Score 55; DB 1; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWGQPHGG 8  
|||||  
DB 56 GWGQPHGG 63

RESULT 14  
US-08-692-892-2  
Sequence 2, Application US/08692892  
Patent No. 5792901  
GENERAL INFORMATION:  
APPLICANT: Prusiner, Stanley B.  
APPLICANT: Scott, Michael R.  
APPLICANT: Telling, Glenn  
TITLE OF INVENTION: DETECTING PRIONS IN A SAMPLE AND  
TITLE OF INVENTION: PRION PREPARATION AND TRANSGENIC ANIMAL USED FOR SAME  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Karl Bozicevic  
STREET: 2200 Sand Hill Road  
CITY: Menlo Park  
STATE: CA  
COUNTRY: USA  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/692,892  
FILING DATE: 30-JULY-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bozicevic, Karl  
REGISTRATION NUMBER: 28, 807  
REFERENCE/DOCKET NUMBER: 06510/056001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 322-5070  
TELEFAX: (415) 854-0875  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 253 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: HUMAN PRION PROTEIN, HuPrP  
US-08-692-892-2

Query Match 100.0%; Score 55; DB 1; Length 253;

Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8  
Db 56 GWGQPHGG 63

## RESULT 15

US-08-713-939A-2  
; Sequence 2, Application US/08713939A  
; Patent No. 5846533  
; GENERAL INFORMATION:  
; APPLICANT: Prusiner, Stanley B.  
; APPLICANT: Williamson, R. Anthony  
; APPLICANT: Burton, Dennis R.  
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP  
; NUMBER OF SEQUENCES: 86  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 2200 Sand Hill Road  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/713,939A  
; FILING DATE: 13-SEP-1996  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bozicevic, Karl  
; REGISTRATION NUMBER: 28,807  
; REFERENCE/DOCKET NUMBER: 06510/059001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-854-5277  
; TELEFAX: 415-854-0875  
; TELEX:  
; INFORMATION FOR SEQ ID NO. 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 253 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-713-939A-2

Query Match 100.0%; Score 55; DB 2; Length 253;  
Best Local Similarity 100.0%; Pred. No. 0.22;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8  
Db 56 GWGQPHGG 63

Search completed: January 3, 2003, 15:34:26  
Job time : 12 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:32:11 : Search time 6.36364 Seconds  
(without alignments)  
23.825 Million cell updates/sec

Title: US-09-543-188a-1  
Perfect score: 55  
Sequence: 1 GWCOPHGG 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 117078 seqs, 18951520 residues

Minimum number of hits satisfying chosen parameters: 117078

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published\_Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	43	9	US-09-939-780-2
2	55	100.0	43	10	US-09-147-761-2
3	55	100.0	253	10	US-09-904-987-3
4	55	100.0	253	10	US-09-919-172-57
5	55	100.0	253	10	US-09-943-906-2
6	55	100.0	254	9	US-10-106-574-5
7	55	100.0	254	9	US-10-106-574-6
8	55	100.0	254	9	US-10-106-574-7
9	55	100.0	254	9	US-10-106-574-8
10	55	100.0	254	10	US-09-943-906-1
11	55	100.0	255	10	US-09-943-906-4
12	55	100.0	256	9	US-10-109-551-2
13	55	100.0	256	9	US-10-109-551-4
14	55	100.0	256	9	US-10-109-551-6
15	55	100.0	256	9	US-10-109-551-8
16	55	100.0	256	9	US-10-109-551-10
17	55	100.0	263	10	US-09-943-906-3
18	55	100.0	439	9	US-10-115-984-2
19	46	83.6	8	12	US-10-035-598-1

20	40.5	73.6	387	10	US-09-866-987-9	Sequence 9, Appl1
21	40	72.7	44	8	US-08-424-550B-519	Sequence 519, App
22	40	72.7	774	9	US-09-945-901-42	Sequence 42, Appl
23	40	72.7	774	9	US-10-007-747-42	Sequence 42, Appl
24	40	72.7	1214	9	US-09-945-901-54	Sequence 54, Appl
25	40	72.7	1214	9	US-10-007-747-54	Sequence 54, Appl
26	40	72.7	1219	9	US-09-945-901-50	Sequence 50, Appl
27	40	72.7	1219	9	US-10-007-747-50	Sequence 50, Appl
28	40	72.7	1231	9	US-09-945-901-48	Sequence 48, Appl
29	40	72.7	1231	9	US-10-007-747-48	Sequence 48, Appl
30	40	72.7	1236	9	US-09-945-901-6	Sequence 6, Appl1
31	40	72.7	1236	9	US-10-007-747-6	Sequence 6, Appl1
32	40	72.7	1239	9	US-09-945-901-52	Sequence 52, Appl
33	40	72.7	1239	9	US-10-007-747-52	Sequence 52, Appl
34	40	72.7	1244	9	US-09-945-901-46	Sequence 46, Appl
35	40	72.7	1244	9	US-10-007-747-46	Sequence 46, Appl
36	38	69.1	5179	9	US-10-025-380-1068	Sequence 1068, Ap
37	38	69.1	5179	10	US-09-922-217-1068	Sequence 1068, Ap
38	38	69.1	5179	10	US-09-833-263-1068	Sequence 1068, Ap
39	37	67.3	40	8	US-08-424-550B-552	Sequence 552, App
40	37	67.3	100	9	US-10-001-835-172	Sequence 172, App
41	37	67.3	242	10	US-09-764-853-609	Sequence 609, App
42	37	67.3	242	10	US-09-764-898-199	Sequence 199, App
43	37	67.3	677	10	US-09-006-298-21	Sequence 21, Appl
44	37	67.3	696	10	US-09-925-300-1253	Sequence 1253, Ap
45	36.5	66.4	67	10	US-09-864-761-44423	Sequence 44423, A

#### ALIGNMENTS

RESULT 1  
US-09-939-780-2  
Sequence 2, Application US/09939780  
Patent No. US20020168689A1  
GENERAL INFORMATION:  
APPLICANT: O'Connor, Michael  
TITLE OF INVENTION: Immunological Assay for Spongiform Encephalopathies  
FILE REFERENCE: 500020US  
CURRENT APPLICATION NUMBER: US/09/939,780  
CURRENT FILING DATE: 2001-08-28  
PRIOR APPLICATION NUMBER: 09/147,761  
PRIOR FILING DATE: 1999-03-03  
PRIOR APPLICATION NUMBER: PCT/IE98/00007  
PRIOR FILING DATE: 1998-02-06  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 2  
LENGTH: 43  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: prion protein  
US-09-939-780-2  
Query Match  
Best Local Similarity 100.0%; Score 55; DB 9; Length 43;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GWCOPHGG 8  
DB 15 GWCOPHGG 22  
RESULT 2  
US-09-147-761-2  
Sequence 2, Application US/09147761  
Patent No. US20010010918A1  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: IMMUNOLOGICAL ASSAY FOR SPONGIFORM  
ENCEPHALOPATHIES

```

; NUMBER OF SEQUENCES: 4
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,761
; FILING DATE:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: WO IE/98/00007
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CHRISTINA GATES
; REFERENCE/DOCKET NUMBER: PL678pct
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 353-1-6605033
; TELEFAX: 353-1-6606920
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: amino acid
; HYPOTHEetical:
; ANTI-SENSE:
; ORIGINAL SOURCE:
; ORGANISM:
; CELL TYPE:
;
US-09-147-761-2

```

Query Match 100.0%; Score 55; DB 10; Length 43;

Best Local Similarity 100.0%; Pred. No. 0.044; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWCQPHGG 8  
DB 15 GWCQPHGG 22

```

RESULT 3
US-09-904-987-3
; Sequence 3, Application US/09904987
; Patent No. US20020037908A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020037908A1acetyl, Inc.
; TITLE OF INVENTION: Methods and Compositions for Controlling Pathological and Prepara
; FILE REFERENCE: 42108/26146
; CURRENT FILING DATE: 2001-07-12
; CURRENT APPLICATION NUMBER: US/09/904,987
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 3
; LENGTH: 253
; TYPE: PRT
; ORGANISM: homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NCBI ENTREZ / XM_009567
; DATABASE ENTRY DATE: 2001-04-17
; RELEVANT RESIDUES: (1)..(253)
US-09-904-987-3

```

Query Match 100.0%; Score 55; DB 10; Length 253;

Best Local Similarity 100.0%; Pred. No. 0.21; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWCQPHGG 8  
DB 56 GWCQPHGG 63

RESULT 4

```

US-09-919-172-57
; Sequence 57, Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIORITY APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 57
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020119463A1 1256895CD1
;
US-09-919-172-57

```

Query Match 100.0%; Score 55; DB 10; Length 253;

Best Local Similarity 100.0%; Pred. No. 0.21; Mismatches 0; Indels 0; Gaps 0;

QY 1 GWCQPHGG 8  
DB 56 GWCQPHGG 63

```

RESULT 5
US-09-943-906-2
; Sequence 2, Application US/09943906
; Patent No. US20020150571A1
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; Williamson, R. Anthony
; Burton, Dennis R.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/943,906
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/550,374
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bozicevic, Karl
; REGISTRATION NUMBER: 28,807
; REFERENCE/DOCKET NUMBER: 06510/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-5277
; TELEFAX: 415-854-0875
; TELEX: <unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-943-906-2
Query Match          100.0%; Score 55; DB 10; Length 253;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8
    |||||||
Db 56 GWGQPHGG 63

RESULT 6
US-10-106-574-5
; Sequence 5, Application US/10106574
; Patent No. US20020164335A1
; GENERAL INFORMATION:
; APPLICANT: Harris, David A.
; TITLE OF INVENTION: Compositions and Methods for the Study and Diagnosis of Prion Dis
; FILE REFERENCE: 09789280.0003
; CURRENT APPLICATION NUMBER: US/10/106,574
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-106-574-5
Query Match          100.0%; Score 55; DB 9; Length 254;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8
    |||||||
Db 63 GWGQPHGG 70

RESULT 7
US-10-106-574-6
; Sequence 6, Application US/10106574
; Patent No. US20020164335A1
; GENERAL INFORMATION:
; APPLICANT: Harris, David A.
; TITLE OF INVENTION: Compositions and Methods for the Study and Diagnosis of Prion Dis
; FILE REFERENCE: 09789280.0003
; CURRENT APPLICATION NUMBER: US/10/106,574
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-106-574-6
Query Match          100.0%; Score 55; DB 9; Length 254;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8
    |||||||
Db 63 GWGQPHGG 70

RESULT 8
US-10-106-574-7
```

```

; Sequence 7, Application US/10106574
; Patent No. US20020164335A1
; GENERAL INFORMATION:
; APPLICANT: Harris, David A.
; TITLE OF INVENTION: Compositions and Methods for the Study and Diagnosis of Prion
; FILE REFERENCE: 09789280.0003
; CURRENT APPLICATION NUMBER: US/10/106,574
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-106-574-7
Query Match          100.0%; Score 55; DB 9; Length 254;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8
    |||||||
Db 63 GWGQPHGG 70

RESULT 9
US-10-106-574-8
; Sequence 8, Application US/10106574
; Patent No. US20020164335A1
; GENERAL INFORMATION:
; APPLICANT: Harris, David A.
; TITLE OF INVENTION: Compositions and Methods for the Study and Diagnosis of Prion
; FILE REFERENCE: 09789280.0003
; CURRENT APPLICATION NUMBER: US/10/106,574
; CURRENT FILING DATE: 2002-03-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 8
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-106-574-8
Query Match          100.0%; Score 55; DB 9; Length 254;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGQPHGG 8
    |||||||
Db 63 GWGQPHGG 70

RESULT 10
US-09-943-906-1
; Sequence 1, Application US/09943906
; Patent No. US20020150571A1
; GENERAL INFORMATION:
; APPLICANT: Prusiner, Stanley B.
; TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```

COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/943,906  
FILING DATE: 30-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/550,374  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Bozicevic, Karl  
REGISTRATION NUMBER: 28,807  
REFERENCE/DOCKET NUMBER: 06510/059001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-854-5277  
TELEFAX: 415-854-0875  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 254 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-943-906-1

Query Match 100.0%; Score 55; DB 10; Length 254;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGPHHG 8  
Db 63 GWGPHHG 70

RESULT 11  
US-09-943-906-4  
Sequence 4, Application US/09943906  
Patent No. US20020150571A1  
GENERAL INFORMATION:  
APPLICANT: Prusiner, Stanley B.  
Williamson, R. Anthony  
Burton, Dennis R.  
TITLE OF INVENTION: ANTIBODIES SPECIFIC FOR NATIVE PrP  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 2200 Sand Hill Road  
CITY: Menlo Park  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 94025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/943,906  
FILING DATE: 30-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/550,374  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Bozicevic, Karl  
REGISTRATION NUMBER: 28,807  
REFERENCE/DOCKET NUMBER: 06510/059001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-854-5277  
TELEFAX: 415-854-0875

TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 255 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-943-906-4

Query Match 100.0%; Score 55; DB 10; Length 255;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGPHHG 8  
Db 58 GWGPHHG 65

RESULT 12  
US-10-109-551-2  
Sequence 2, Application US/10109551  
Publication No. US20020194635A1  
GENERAL INFORMATION:  
APPLICANT: DUNNE, PATRICK W.  
TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE  
FILE REFERENCE: TAMK:20705  
CURRENT APPLICATION NUMBER: US/10/109,551  
CURRENT FILING DATE: 2002-03-28  
PRIOR APPLICATION NUMBER: 60/280,549  
PRIOR FILING DATE: 2001-03-30  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 256  
TYPE: PrT  
ORGANISM: Bos taurus  
US-10-109-551-2

Query Match 100.0%; Score 55; DB 9; Length 256;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGPHHG 8  
Db 59 GWGPHHG 66

RESULT 13  
US-10-109-551-4  
Sequence 4, Application US/10109551  
Publication No. US20020194635A1  
GENERAL INFORMATION:  
APPLICANT: DUNNE, PATRICK W.  
TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE  
FILE REFERENCE: TAMK:20705  
CURRENT APPLICATION NUMBER: US/10/109,551  
CURRENT FILING DATE: 2002-03-28  
PRIOR APPLICATION NUMBER: 60/280,549  
PRIOR FILING DATE: 2001-03-30  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 256  
TYPE: PrT  
ORGANISM: Ovis aries  
US-10-109-551-4



Query Match 100.0%; Score 55; DB 9; Length 256;  
 Best Local Similarity 100.0%; Pred. No. 0.21;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGPHGG 8  
 |||||  
 Db 59 GWGPHGG 66

RESULT 14  
 US-10-109-551-6  
 ; Sequence 6, Application US/10109551  
 ; Publication No. US20020194635A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DUNNE, PATRICK W.  
 ; APPLICANT: PIEDRAHITA, JORGE  
 ; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE  
 ; FILE REFERENCE: TARK:207US  
 ; CURRENT APPLICATION NUMBER: US/10/109,551  
 ; PRIOR FILING DATE: 2002-03-28  
 ; PRIOR APPLICATION NUMBER: 60/280,549  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 256  
 ; TYPE: PRT  
 ; ORGANISM: Odocolleus virginianus  
 US-10-109-551-6

Query Match 100.0%; Score 55; DB 9; Length 256;  
 Best Local Similarity 100.0%; Pred. No. 0.21;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGPHGG 8  
 |||||  
 Db 59 GWGPHGG 66

RESULT 15  
 US-10-109-551-8  
 ; Sequence 8, Application US/10109551  
 ; Publication No. US20020194635A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DUNNE, PATRICK W.  
 ; APPLICANT: PIEDRAHITA, JORGE  
 ; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE  
 ; FILE REFERENCE: TARK:207US  
 ; CURRENT APPLICATION NUMBER: US/10/109,551  
 ; CURRENT FILING DATE: 2002-03-28  
 ; PRIOR APPLICATION NUMBER: 60/280,549  
 ; PRIOR FILING DATE: 2001-03-30  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 8  
 ; LENGTH: 256  
 ; TYPE: PRT  
 ; ORGANISM: Odocolleus hemionus hemionus  
 US-10-109-551-8

Query Match 100.0%; Score 55; DB 9; Length 256;  
 Best Local Similarity 100.0%; Pred. No. 0.21;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GWGPHGG 8  
 |||||  
 Db 59 GWGPHGG 66

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:21:45 : Search time 26.5909 seconds  
(without alignments)  
30.067 Million cell updates/sec

Title: US-09-543-188a-3  
Perfect score: 34  
Sequence: 1 LLIWIP 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq\_101002:\*

1:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA2001.DAT:*
23:	/SIDS2/gcgdata/geneseq/genesep-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	34	100.0	6	23	AAU11826	Peptide ligand for
2	34	100.0	87	22	ABB40848	Peptide #8354 enco
3	34	100.0	87	22	ABB25012	Protein #7011 enco
4	34	100.0	87	22	AA661707	Human brain expres
5	34	100.0	87	22	AA674503	Human bone marrow
6	34	100.0	87	22	AA634616	Peptide #8653 enco
7	34	100.0	87	23	ABG44369	Human peptide enco
8	34	100.0	269	21	ABB26446	Drosophila melanog
9	34	100.0	269	22	ABB68723	Drosophila melanog
10	34	100.0	453	21	AA620922	Drosophila odorant

11	34	100.0	554	22	AAU14133	Human novel protei
12	34	100.0	754	23	ABB92828	Herbicide1dly activ
13	33	97.1	298	20	AAV24023	Human CMRF-35-H9 r
14	33	97.1	301	20	AAV24022	Human CMRF-35-H9 r
15	32	94.1	128	15	AA653332	KM-603 light chain
16	32	94.1	128	20	AAV28360	Antibody light chain
17	32	94.1	142	16	AA681311	Rat monoclonal ant
18	32	94.1	177	22	AA615989	Human novel secret
19	32	94.1	251	22	AAU33248	Novel human secret
20	32	94.1	484	22	ABG05597	Novel human diago
21	31	91.2	12	19	AAW60537	Oligopeptide from
22	31	91.2	16	17	AA684070	Plasmod blue-Tp6
23	31	91.2	20	12	AA610540	Eukaryotic signal
24	31	91.2	20	19	AAW69959	Ig kappa chain lea
25	31	91.2	20	19	AAW68295	Ig kappa chain lea
26	31	91.2	20	20	AAV17756	Target signal pept
27	31	91.2	20	21	AAV52559	Consensus murine I
28	31	91.2	20	22	AAE11049	Human Ig (immunog
29	31	91.2	20	22	AAE08740	Immunoglobulin kap
30	31	91.2	20	22	AAE04428	Human immunoglobul
31	31	91.2	20	22	AAE01995	Immunoglobulin (Ig
32	31	91.2	21	18	AAW12384	Murine Ig kappa-ch
33	31	91.2	21	22	AA686331	Ig-kappa chain lea
34	31	91.2	21	23	AAU75372	Mouse kappa immuno
35	31	91.2	21	23	AAU75379	Mouse kappa immuno
36	31	91.2	42	22	AA671931	Murine sig-mend f
37	31	91.2	45	21	AA603863	Active human neuro
38	31	91.2	47	23	AAW47704	Fragment of a reco
39	31	91.2	65	21	AA624836	Plant SDF encoded
40	31	91.2	69	22	AAU21012	Human novel foetal
41	31	91.2	80	21	AAV52548	Murine Ig-kappa si
42	31	91.2	91	22	AAU14672	Novel bone marrow
43	31	91.2	100	21	AAV64671	Human 5' EST relat
44	31	91.2	116	22	AAU44203	Protonlactetium
45	31	91.2	118	21	AAV52541	KapplaMP-Th fusio

## ALIGNMENTS

RESULT 1	
AAU11826	AAU11826 standard; peptide: 6 AA.
ID	XX
AC	AAU11826;
XX	XX
DT	26-MAR-2002 (first entry)
XX	XX
DE	Peptide ligand for Prion protein, Prp, #1.
XX	XX
KW	Prion protein; Prp; ligand; octapeptide motif; scrapie;
KW	prion-associated disease; Creutzfeldt-Jakob disease;
KW	Gerstmann-Strausler-Scheinker disease; fatal familial insomnia;
KW	feline spongiform encephalopathy; bovine spongiform encephalopathy;
KW	transmissible mink encephalopathy; exotic ungulate encephalopathy;
OS	chronic wasting disease.
XX	XX
XX	Synthetic.
XX	XX
PN	WO200177687-A2.
XX	XX
PD	18-OCT-2001.
XX	XX
PF	05-APR-2001; 2001WO-US11150.
XX	XX
PR	05-APR-2000; 2000US-0543188.
XX	XX
PA	(VITE-) VI TECHNOLOGIES INC.
XX	XX
PI	Hammond DJ, Wiltshire VR, Carbonell R, Shen H;
XX	XX
DR	WPI; 2002-061944/08.
XX	XX

PT New ligands for prion proteins, useful for detection or removal of  
PT prions and for treating prion-associated diseases, recognize a specific  
PT octapeptide motif -  
XX  
PS Claim 16: Page 34, 47pp; English.

XX The invention relates to a ligand of less than 6 kD that binds to a  
CC polypeptide containing the sequence GlyTyrGlyGlnProHisGly (A) or an  
CC analogue that is the retro-inverso isomer of (A). The sequence A is  
CC an octapeptide motif from the prion protein (PrP). The ligands are  
CC identified by binding assays with the peptide (A) or peptides containing  
CC (A). The ligands are used for detecting prion proteins (or prions) in  
CC biological or environmental samples, e.g. for diagnosis, also for  
CC removing them from samples, for treating or retarding development of  
CC prion-associated diseases, especially Creutzfeldt-Jakob diseases (in  
CC latrogenic, new variant, familial or sporadic forms), but also  
CC Gerstmann-Strausser-Scheinker disease, fatal familial insomnia,  
CC scrapie, bovine or feline spongiform encephalopathy, transmissible mink  
CC or exotic ungulate encephalopathy, or chronic wasting disease. The  
CC present sequence is a ligand of the invention.

Sequence 6 AA:

Query Match 100.0%; Score 34; DB 23; Length 6;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
1 |||||  
DB 1 LLIWIP 6

RESULT 2  
ABB40848

ID ABB40848 standard; Peptide: 87 AA.

XX ABB40848;

DT 04-FEB-2002 (first entry)

XX Peptide #8354 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

PE 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human foetal liver -

PS Claim 27; SEQ ID NO 33483; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC foetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [ftp:wipo.int/pub/published\\_pct\\_sequences](ftp:wipo.int/pub/published_pct_sequences).

Sequence 87 AA:

Query Match 100.0%; Score 34; DB 22; Length 87;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
48 |||||  
DB 48 LLIWIP 53

RESULT 3  
ABB25012

ID ABB25012 standard; Protein: 87 AA.

XX ABB25012;

DT 23-JAN-2002 (first entry)

XX Protein #7011 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;

XX cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

PE 30-JAN-2001; 2001WO-US00666.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts -

PS Claim 15; SEQ ID NO 26782; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [ftp:wipo.int/pub/published\\_pct\\_sequences](ftp:wipo.int/pub/published_pct_sequences).

SQ Sequence 87 AA;

Query Match 100.0%; Score 34; DB 22; Length 87;

Best Local Similarity 100.0%; Pred. No. 54; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
|||||  
Db 48 LLIWIP 53

RESULT 4  
AAM61707  
ID AAM61707 standard; Protein; 87 AA.

AC AAM61707;

DT 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe encoded protein SEQ ID NO: 33812.

XX Human; brain expressed exon; gene expression analysis; probe;  
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
XX epilepsy; cancer.

OS Homo sapiens.

PN WO200157275-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human  
XX brains -

PS Example 4; SEQ ID NO: 33812; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX brain. They can be used to measure gene expression in brain cell samples,  
XX which may enable the diagnosis and improved treatment of nervous system  
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
XX epilepsy and cancers. The present sequence is a protein encoded by one of  
XX the probes of the invention.

XX Sequence 87 AA;

Query Match 100.0%; Score 34; DB 22; Length 87;

Best Local Similarity 100.0%; Pred. No. 54; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
|||||  
Db 48 LLIWIP 53

RESULT 5  
AAM74503

ID AAM74503 standard; Protein; 87 AA.

XX AAM74503;

XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 34809.  
XX Human; bone marrow expressed exon; gene expression analysis; probe;  
XX microarray; cancer; leukaemia; lymphoma; myeloma.

OS Homo sapiens.

PN WO200157276-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human bone marrow -

PS Example 4; SEQ ID NO: 34809; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX bone marrow. They can be used to measure gene expression in bone marrow  
XX samples, which may enable the improved diagnosis and treatment of cancers  
XX such as lymphoma, leukaemia and myeloma. The present sequence is a  
XX protein encoded by one of the probes of the invention.

XX Sequence 87 AA;

Query Match 100.0%; Score 34; DB 22; Length 87;

Best Local Similarity 100.0%; Pred. No. 54; Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
|||||  
Db 48 LLIWIP 53

RESULT 6  
AAM34616  
ID AAM34616 standard; Protein; 87 AA.

XX AAM34616;

XX 17-OCT-2001 (first entry)

XX Peptide #8653 encoded by probe for measuring placental gene expression.

XX Probe: microarray; human; placenta; antenatal diagnosis;  
XX genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

PD 09-AUG-2001.  
XX  
XX 30-JAN-2001: 2001MO-US00663.  
XX  
XX 04-FEB-2000: 2000US-0180312.  
PR 26-MAY-2000: 2000US-0207456.  
PR 30-JUN-2000: 2000US-0608408.  
PR 03-AUG-2000: 2000US-0632366.  
PR 21-SEP-2000: 2000US-0234687.  
PR 27-SEP-2000: 2000US-0236359.  
PR 04-OCT-2000: 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR:  
XX WPI: 2001-48897/53.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
XX Claim 27: SEQ ID NO 34885: 654pp: English.  
XX  
XX The present invention relates to single exon nucleic acid probes (SENP:  
CC see AAI31315-AA157546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.  
XX  
XX Sequence 87 AA:  
SQ  
Query Match 100.0%; Score 34; DB 22; Length 87;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLIWIP 6  
|||||  
Db 48 LLIWIP 53  
RESULT 7  
ID ABB44369 standard; Peptide: 87 AA.  
XX  
XX ABB44369;  
AC  
XX 19-AUG-2002 (first entry)  
PT  
XX Human peptide encoded by genome-derived single exon probe SEQ ID 34034.  
XX  
XX Human: single exon probe; asthma; lung cancer; COPD; ILD;  
KM chronic obstructive pulmonary disease; interstitial lung disease;  
KM familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
KM Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
KM primary ciliary dyskinesia; pulmonary hypertension;  
KM hyaline membrane disease.  
XX  
XX Homo sapiens.  
XX  
XX WO200186003-A2.  
PN  
XX 15-NOV-2001.  
PD  
XX 30-JAN-2001: 2001MO-US00665.  
PE  
XX 04-FEB-2000: 2000US-180312P.  
PR 26-MAY-2000: 2000US-207456P.  
PR 30-JUN-2000: 2000US-0608408.  
PR 03-AUG-2000: 2000US-0632366.

PR 21-SEP-2000: 2000US-234687P.  
PR 27-SEP-2000: 2000US-236359P.  
PR 04-OCT-2000: 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR:  
XX WPI: 2002-114183/15.  
XX  
XX Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples -  
XX  
XX Claim 27: SEQ ID NO 34034: 634pp: English.  
XX  
XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of  
CC probes; the novel set of probes which hybridize at high stringency to a  
CC nucleic acid expressed in the human lung; measuring gene expression in a  
CC sample derived from human lung, comprising (a) contacting the array with  
CC a collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of  
CC the array; identifying exons in a eukaryotic genome, comprising  
CC (a) algorithmically predicting at least one exon from genomic sequences  
CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene  
CC expression analysis, and for identifying exons in a gene, particularly  
CC using human lung derived mRNA and for the study of lung diseases  
CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic  
CC and hyaline membrane disease. The present sequence is a peptide/protein  
CC encoded by a single exon probe of the invention.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
XX Sequence 87 AA:  
SQ  
Query Match 100.0%; Score 34; DB 23; Length 87;  
Best Local Similarity 100.0%; Pred. No. 54;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLIWIP 6  
|||||  
Db 48 LLIWIP 53  
RESULT 8  
ID AAB26446 standard; Protein: 269 AA.  
XX  
XX AAB26446;  
AC  
XX

DT 23-FEB-2001 (first entry)  
XX  
PA Drosophila melanogaster odorant receptor DOR59.  
DE  
XX  
KW Odorant receptor; fruit fly; DOR59; odour recognition; pest control.  
XX  
XX Drosophila melanogaster.  
OS  
PN WO200050566-A2.  
XX  
PD 31-AUG-2000.  
XX  
XX 25-FEB-2000; 2000WO-US04995.  
PF  
XX 25-FEB-1999; 99US-0257706.  
PR  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
PA  
XX Vossball LB, Amrein HO, Axel R;  
PI  
XX WPI: 2000-572081/53.  
XX N-PSDB; AAA94845.  
PT Novel nucleic acid encoding an insect odorant receptor, for identifying  
XX modulator compounds that are useful in controlling pest population  
XX  
PS Disclosure; Page 75; 176pp; English.  
XX  
XX The present sequence is the previously identified Drosophila melanogaster  
CC odorant receptor DOR59. The odorant genes and proteins, such as those  
CC provided by the invention, are useful as they aid in the study of the  
CC olfactory organ in mammals, as well as aiding the understanding of the  
CC link between odour recognition and behaviour in insects. They also enable  
CC the identification of compounds capable of activating and inhibiting the  
CC receptors, allow the control of pest populations via the use of alarm  
CC odour ligands and via the use of ligands which interfere with the  
CC interaction between odorant ligands and receptors associated with  
CC fertility.  
XX  
SQ Sequence 269 AA:  
Query Match 100.0%; Score 34; DB 21; Length 269;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LTIWIP 6  
DB |||||  
24 LTIWIP 29  
RESULT 9  
AAB68723  
ID ABB68723 standard; Protein; 269 AA.  
XX  
AC ABB68723;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 32961.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
PD  
XX 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
XX  
PR 11-JUL-2000; 2000US-0614150.  
XX

XX  
XX (PEKE ) PE CORP NY.  
PA  
XX  
PI Venter JC, Adams M, Li PMD, Myers EW;  
XX  
DR WPI: 2001-656860/75.  
XX  
DR N-PSDB; ABL12826.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Disclosure; SEQ ID NO 32961; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL101840-ABL16175) and the encoded proteins  
CC (AAB57737-AAB572072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
SQ Sequence 269 AA:  
Query Match 100.0%; Score 34; DB 22; Length 269;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LTIWIP 6  
DB |||||  
24 LTIWIP 29  
RESULT 10  
AAB20922  
ID AAB20922 standard; Protein; 453 AA.  
XX  
AC AAB20922;  
XX  
DT 06-DEC-2000 (first entry)  
XX  
DE Drosophila odorant receptor DOR 56E.1.  
XX  
XX Odorant receptor; Drosophila; olfactory receptor;  
KW G protein-coupled receptor; GPCR superfamily; transgenic insect;  
KW insect behaviour modification; pest control; pollinator attraction;  
KW biosensor; odour detection; odour identification; apiculture.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200043410-A2.  
PD  
XX 27-JUL-2000.  
XX  
PF 25-JAN-2000; 2000WO-US01823.  
XX  
PR 25-JAN-1999; 99US-0117132.  
XX  
PA (UYVA ) UNIV YALE.  
XX  
XX Carlson JR, Kim J, Clyne PJ, Warr CG;  
XX  
DR WPI: 2000-543246/49.  
XX  
DR N-PSDB; AAA72257.  
XX  
PT New nucleic acid encoding a Drosophila olfactory receptor, useful for  
PT identifying modulating agents -  
XX  
PS Claim 12; Page 175-177; 303pp; English.  
XX

XX Sequences AAB20901-B20949 represent *Drosophila melanogaster* odorant  
CC receptors. These proteins function as olfactory receptors, and  
CC are thought to be members of the G protein-coupled receptor  
CC (GPCR) superfamily, which is characterised by the presence of 7  
CC transmembrane helices. Nucleic acids encoding the *Drosophila* odorant  
CC receptors may be used to generate expression constructs, host cells  
CC containing such constructs, and transgenic insects. Cells which express  
CC the odorant receptor genes may be used in methods to identify agents  
CC which modulate expression of these genes, and in methods to identify  
CC receptor binding partners. The *Drosophila* odorant receptor nucleic acids  
CC may also be used to identify corresponding genes in other insects,  
CC such as those which damage crops or transmit disease. The odorant  
CC receptor proteins may be used to identify agents which modulate their  
CC activity, to identify binding partners, as antigens to raise antibodies,  
CC and in methods to modify insect behaviour. The proteins may be also  
CC be used in methods of behaviour modification. Such methods may be used  
CC to study or modify insect behaviour in response to odorants such as  
CC pheromones. Modification of insect behaviour has a wide range of  
CC applications, such as in pest control (e.g., by disrupting the feeding  
CC or mating behaviours of pest species), or for enhancing plant  
CC pollination (by attracting pollinator species). Odorant receptor proteins  
CC and/or nucleotides may also be used to identify appetite suppressants, to  
CC trap odours of a specific type, as biosensors for the detection of  
CC explosives, drugs, perfumes or pollutants, and in apiculture to modify  
CC the behaviour of bees, for example, to increase the production of royal  
CC jelly.

SO Sequence 453 AA:

Query Match 100.0%; Score 34; DB 21; Length 453;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LTIWIP 6  
| | | | |  
DB 159 LTIWIP 164

RESULT 11

AAU14133  
ID AAU14133 standard; Protein: 554 AA.

AC AAU14133;

DT 24-OCT-2001 (first entry)

XX Human novel protein #4.

Human: novel protein; Antianemic; osteopathic; antiinflammatory;  
KW immunomodulatory; cytostatic; neuroprotective; vulnary; nocrotic;  
KW anticonvulsant; antiallergic; cerebroprotective; antifungal; antiviral;  
KW antibacterial; antiallergic; dermatological; haemostatic; antiaslatic;  
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
KW tissue regeneration; immune disorder.

OS Homo sapiens.

PN WO200155437-A2.

PD 02-AUG-2001.

PF 25-JAN-2001; 2001WO-US02623.

PR 25-JAN-2000; 2000US-0491404.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-451939/48.

DR N-PSDB; JAS22438.

XX Isolated polypeptides useful for treating anti-inflammatory diseases,  
PT nervous system disorders, and for regenerating bone and cartilage -  
XX Example 4; Page 522-523; 894pp; English.

XX The invention relates to polynucleotides encoding novel human  
CC proteins or their active domains. The polypeptides, polynucleotides and  
CC antibodies raised against the polypeptides are used in a method of  
CC treatment of a mammal and prevention of disorders caused by the aberrant  
CC protein expression or activity. The polypeptides can be used as  
CC molecular weight markers, food supplements, and in antibody production.  
CC The polypeptides are used to identify compounds which bind to the  
CC polypeptides. Polynucleotides of the invention are used as probes and  
CC primers, for sequencing, for chromosome or gene mapping, in the  
CC production of recombinant proteins, and in generating anti-sense DNA or  
CC RNA and in gene therapy. Polypeptides of the invention can be used to  
CC target drugs to a tumour, in assays to determine biological activity, to  
CC raise antibodies/elicite an immune response, to determine quantitative  
CC protein levels, as tissue markers, and to isolate receptors or ligands.  
CC Polypeptides of the invention may also be useful in treating platelet  
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
CC the proliferation, differentiation and survival of stem cells, as a  
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
CC fungal infection or from autoimmunity, cancer, allergy, asthma,  
CC graft-versus-host disease, eczema, haemophilia, thrombosis,  
CC anti-inflammatory diseases, nervous system disorders, and infection.  
CC The present sequence represents a protein of the invention.

SO Sequence 554 AA:

Query Match 100.0%; Score 34; DB 22; Length 554;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LTIWIP 6  
| | | | |  
DB 138 LTIWIP 143

RESULT 12

ABB92828  
ID ABB92828 standard; Protein: 754 AA.

AC ABB92828;

DT 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 2039.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

OS WO200210210-A2.

PN 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EF09892.

PR 28-AUG-2001; 2001WO-EF09892.

PA (FARB ) BAYER AG.

PI Tietjen K, Weidler M;

XX WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,  
comprising aligning and comparing nucleic acid or amino acid sequences



PT from plant with nucleic acid or amino acid sequences from non-plant  
XX organisms -  
XX  
PS Claim 5; SEQ ID NO 2039; 261pp + Sequence Listing; English.  
XX  
CC The invention relates to identifying target proteins  
CC (AB90790-AB94016) for herbicidally active compounds, comprising  
CC aligning and comparing nucleic acid or amino acid sequences from plant  
CC with nucleic acid or amino acid sequences from non-plant organisms using  
CC suitable search parameters, where plant sequences having an E-value  
CC greater by a factor of 3 than the E-value of most similar non-plant  
CC sequences are selected. The polypeptides or nucleic acids encoding them  
CC are useful for identifying modulators. The identified modulators are  
CC useful as herbicides.  
XX  
SQ Sequence 754 AA:  
XX  
Query Match 100.0%; Score 34; DB 23; Length 754;  
Best Local Similarity 100.0%; Pred. No. 4.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
1 LTIWIP 6  
19 LTIWIP 24  
XX  
Db  
XX  
RESULT 13  
AAAY24023  
ID AAAY24023 standard; Protein; 298 AA.  
XX  
AC AAAY24023;  
XX  
DT 30-SEP-1999 (first entry)  
XX  
DE Human CMRF-35-H9 receptor protein.  
XX  
XX CMRF-35-H9 receptor; immunoglobulin M; IgM; immunomodulation;  
XX humoral immune response; transplant; myeloid leukemia;  
XX promyelocytic leukemia; leukemia.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 10  
FT /note="encoded by CTC"  
XX  
PN WO936438-A1.  
XX  
22-JUL-1999.  
XX  
PF 14-JAN-1999; 99MO-NZ00003.  
XX  
PR 14-JAN-1998; 98NZ-0329582.  
XX  
XX (ORDE-) ORDER OF SISTERS OF MERCY IN QUEENSLAND.  
XX  
PI Hart DNF;  
XX  
XX WPI: 1999-458446/38.  
XX  
XX N-PSDB; AAX86359.  
XX  
XX Human immunoglobulin M receptor CMRF-35-H9, useful for  
XX immunomodulation  
XX  
XX Claim 3; Page 33; 40pp; English.  
XX  
XX The present sequence represents a novel receptor, CMRF-35-H9, which binds  
XX immunoglobulin M (IgM). CMRF-35-H9, a CMRF-35-H9 extracellular domain  
XX peptide or a vector encoding these, can be used to modulate an immune  
XX response in a patient or to block or inhibit a humoral immune response  
XX in a patient. This is particularly useful in a patient who has or is  
XX about to receive a transplant. The CMRF-35-H9 receptor or extracellular  
XX domain binds to antibodies, which are then no longer free to bind to the

CC transplant. A ligand/antigen construct that binds CMRF-35-H9 can be used  
CC to load a protective antigen or a stimulatory antigen into an antigen  
CC presenting cell. The level of CMRF-35-H9 in a patient sample can be  
CC determined to diagnose myeloid leukemia in a patient. An increased level  
CC of CMRF-35-H9 or decreased level of promyelocytic leukemia is diagnostic  
CC of leukemia.  
XX  
SQ Sequence 298 AA:  
XX  
Query Match 97.1%; Score 33; DB 20; Length 298;  
Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
XX  
1 LTIWIP 6  
8 LTIWIP 13  
XX  
Db  
XX  
RESULT 14  
AAAY24022  
ID AAAY24022 standard; Protein; 301 AA.  
XX  
AC AAAY24022;  
XX  
DT 30-SEP-1999 (first entry)  
XX  
XX Human CMRF-35-H9 receptor protein.  
XX  
XX CMRF-35-H9 receptor; immunoglobulin M; IgM; immunomodulation;  
XX humoral immune response; transplant; myeloid leukemia;  
XX promyelocytic leukemia; leukemia.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..13  
FT /label="signal\_peptide"  
FT Misc-difference 10  
FT /note="encoded by CTC"  
FT Protein 14..301  
FT /label="mature\_protein"  
FT Domain 29..126  
FT /note="putative igm binding domain"  
FT Region 178..201  
FT /note="transmembrane region"  
XX  
XX WO936438-A1.  
XX  
22-JUL-1999.  
XX  
PF 14-JAN-1999; 99MO-NZ00003.  
XX  
PR 14-JAN-1998; 98NZ-0329582.  
XX  
XX (ORDE-) ORDER OF SISTERS OF MERCY IN QUEENSLAND.  
XX  
PI Hart DNF;  
XX  
XX WPI: 1999-458446/38.  
XX  
XX N-PSDB; AAX86358.  
XX  
XX Human immunoglobulin M receptor CMRF-35-H9, useful for  
XX immunomodulation  
XX  
XX Claim 12; Fig 2; 40pp; English.  
XX  
XX The present sequence represents a novel receptor, CMRF-35-H9, which binds  
XX immunoglobulin M (IgM). CMRF-35-H9, a CMRF-35-H9 extracellular domain  
XX peptide or a vector encoding these, can be used to modulate an immune  
XX response in a patient or to block or inhibit a humoral immune response  
XX in a patient. This is particularly useful in a patient who has or is  
XX about to receive a transplant. The CMRF-35-H9 receptor or extracellular  
XX domain binds to antibodies, which are then no longer free to bind to the

CC transplant. A ligand/antigen construct that binds CMRF-35-H9 can be used  
CC to load a protective antigen or a stimulatory antigen into an antigen  
CC presenting cell. The level of CMRF-35-H9 in a patient sample can be  
CC determined to diagnose myeloid leukemia in a patient. An increased level  
CC of CMRF-35-H9 or decreased level of promyelocytic leukemia is diagnostic  
CC of leukemia.  
CC  
SQ Sequence 301 AA;  
SQ  
Query Match 97.1%; Score 33; DB 20; Length 301;  
Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLIWIP 6  
DB 8 LLIWIP 13  
RESULT 15  
AAR53332  
AAR53332 standard; Protein; 128 AA.  
AC AAR53332;  
DT 17-NOV-1994 (first entry)  
DE KM-603 light chain.  
XX  
XX Monoclonal antibody; Ab; ganglioside GM2; chimera;  
KW chimeric antibody; expression vector; heavy; light; chain;  
KW hypervariable region; CDR; constant region; hybridoma;  
KW Ig; immunoglobulin; KM-796; KM-750; KM-603; cancer.  
OS  
XX Rat rattus.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..20  
FT /label= sig\_peptide  
FT Region 44..54  
FT /label= CDR1  
FT Region 70..76  
FT /label= CDR2  
FT Region 109..117  
FT /label= CDR3  
XX  
XX AU9346181-A.  
XX 17-MAR-1994.  
PD  
XX 07-SEP-1993; 93AU-0046181.  
XX  
XX 07-SEP-1992; 92JP-0238452.  
XX  
XX (KYOW ) KYOMA HAKKO KOGYO KK.  
XX  
XX Hanai N, Hasegawa M, Kolke M, Kuwana Y, Nakamura K;  
XX Shitara K;  
XX  
XX MPI: 1994-126857/16.  
XX N-PSDB; AAQ45430.  
XX  
XX Humanised antibody specific for ganglioside GM2 - used for  
XX producing a cytotoxic effect on cancers such as melanoma,  
XX neuroblastoma and glioma.  
XX  
XX Disclosure: Page 110; 191pp; English.  
XX  
XX Chimeric human Ab expression vectors are constructed by inserting  
XX the Ab heavy and light chain variable region-encoding CDNA  
XX isolated from hybridomas producing a mouse or rat monoclonal Ab  
XX reacting with the ganglioside GM2 respectively into an expression  
XX vector for use in animal cells which contains the human Ab heavy and  
XX light chain constant region-encoding CDNA. The expression vectors

CC are introduced into animal cells and the transformant thus obtained  
CC is cultured for the prodn. of a chimeric human Ab reacting with the  
CC ganglioside GM2. In contrast to mouse monoclonal Abs, the chimeric  
CC human Abs will not cause anti-mouse Ig Ab prodn. in the patient's  
CC body but show a prolonged blood half-life, with a reduced frequency  
CC of adverse effects, so that it can be expected to be superior  
CC to mouse monoclonal Abs in the efficacy in the treatment of human  
CC cancer, for instance.  
CC Mouse anti-GM2 monoclonal Ab KM-796 and KM-750 and rat KM-603  
CC heavy and light chain sequences are given in AAQ45426-30.  
CC CDR regions for use in chimeric Abs are indicated in the  
CC Features Table.  
XX  
SQ Sequence 128 AA;  
SQ

Query Match 94.1%; Score 32; DB 15; Length 128;  
Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
DB 11 LLIWIP 16

Search completed: January 3, 2003, 15:28:35  
Job time : 27.5909 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:00 ; Search time 9.95455 Seconds  
(without alignments)  
57.944 Million cell updates/sec

Title: US-09-543-188A-3

Perfect score: 34

Sequence: 1 LLIWIP 6

Scoring table: BLASTSUM62

Searched: 283224 seqs, 96134422 residues

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	34	100.0	367	F70036	capsular polysacch
2	34	100.0	402	S65973	transport protein
3	34	100.0	554	S71751	dolichyl-phosphate
4	34	100.0	754	A85043	probable LRR recep
5	34	100.0	2452	RN202L	DNA-directed RNA p
6	34	100.0	2457	T18492	hypothetical prote
7	33	97.1	171	C45665	adult-specific bru
8	33	97.1	304	S69034	probable membrane
9	32	94.1	126	S40312	Ig kappa chain - h
10	32	94.1	200	G83834	hypothetical prote
11	32	94.1	269	T15500	hypothetical prote
12	32	94.1	501	G82822	NADH2 dehydrogenas
13	32	94.1	542	T19862	hypothetical prote
14	31	91.2	114	S00996	Ig kappa chain pre
15	31	91.2	115	S10146	Ig kappa chain pre
16	31	91.2	117	S24207	Ig kappa chain v r
17	31	91.2	120	S06731	Ig kappa chain pre
18	31	91.2	120	S06732	Ig kappa chain pre
19	31	91.2	128	JL0073	aberrant kappa tra
20	31	91.2	131	KVMSM6	Ig kappa chain pre
21	31	91.2	131	PH1226	Ig kappa chain pre
22	31	91.2	131	S55027	Ig light chain pre
23	31	91.2	132	G83872	hypothetical prote
24	31	91.2	131	KVMS32	Ig kappa chain pre
25	31	91.2	140	PNO446	Ig kappa chain pre
26	31	91.2	157	B86424	unknown protein, 3
27	31	91.2	229	T34215	Fc gamma (IgG) rec
28	31	91.2	280	S55577	Fc gamma (IgG) rec
29	31	91.2	336	I48471	Fc gamma (IgG) rec

30	31	91.2	344	A41357	Fc gamma (IgG) rec
31	31	91.2	349	B86603	muramoyl-pentapept
32	31	91.2	349	G72021	muramoyl-pentapept
33	31	91.2	349	A81519	phospho-N-acetylmu
34	31	91.2	365	A48105	probable M-factor
35	31	91.2	374	A39878	Fc gamma (IgG) rec
36	31	91.2	396	B65169	multidrug resistanc
37	31	91.2	396	F91205	2-module integral
38	31	91.2	396	B86051	2-module integral
39	31	91.2	404	A46480	Fc gamma (IgG) rec
40	31	91.2	431	B84069	hypothetical prote
41	31	91.2	443	C71904	anaerobic c4-dicar
42	31	91.2	443	D64610	dicarboxylate memb
43	31	91.2	500	AF2325	NADH dehydrogenase
44	31	91.2	512	T15669	hypothetical prote
45	31	91.2	741	A83271	hypothetical prote

## ALIGNMENTS

RESULT 1  
F70036  
capsular polysaccharide biosynthesis homolog yveQ - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
C:Accession: F70036  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capriano, V.; Carter, N.M.;  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funna, S.; Gallizzi, A.; Gal  
isch, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scani  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiya  
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: F70036  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-367 <KUN>  
A:Cross-references: GB:299121; GB:AL009126; NID:92635827; PIDN:CAB15436.1; PID:el1861  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yveQ  
Query Match  
Best Local Similarity 100.0%; Score 34; DB 2; Length 367;  
Pred. No. 64;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLIWIP 6  
Db 37 LLIWIP 42  
RESULT 2  
S65973  
transport protein homolog yycB - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: S65973; A70089  
R:Ogasawara, N.; Nakai, S.; Yoshikawa, H.  
DNA Res. 1, 1-14, 1994  
A:Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtilis ch  
A:Reference number: S65967; MUID:96051385; PMID:7584024  
A:Accession: S65973  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown



A:Molecule type: DNA  
 A:Residues: 1-2457 <LAW>  
 A:Cross-references: EMBL:Z98551; PIDN:CAB1131.1  
 C:Genetics:  
 A:Map position: 3  
 A>Note: C0805W  
 C:Superfamily: Plasmodium DNA-directed RNA polymerase II large chain

Query Match 100.0%; Score 34; DB 2; Length 2457;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
 |||||  
 DB 582 LLIWIP 587

## RESULT 7

■ nt-specific brush border protein - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 08-Oct-1999  
 C:Accession: C45665  
 R:Boil, W.; Schmid-Chanda, T.; Semenza, G.; Mantel, N.  
 J. Biol. Chem. 268, 12901-12911, 1993  
 A:Title: Messenger RNAs expressed in intestine of adult but not baby rabbits. Isolation

A:Reference number: A45665; MUID:93286138; PMID:8509424

A:Accession: C45665

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-171 <BO>

A:Cross-references: GB:Z12842; NID:q1763; PIDN:CA78304.1; PID:q1764

C:Superfamily: rabbit adult-specific brush border protein

C:Keywords: intestine; transmembrane protein

Query Match 97.1%; Score 33; DB 2; Length 171;  
 Best Local Similarity 83.3%; Pred. No. 47;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
 |||||  
 DB 63 LLIWIP 68

## RESULT 8

■ able membrane protein YPR147c - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 22-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 19-Apr-2002

C:Accession: S69034

R:Fulton, L.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of S. cerevisiae cosmid 9659.

A:Reference number: S69022

A:Accession: S69034

A:Molecule type: DNA

A:Residues: 1-304 <FU>

A:Cross-references: EMBL:U40829; NID:q1066476; PID:q1066494; GSPDB:GN00016; MIPS:YPR147C

C:Genetics:

A:Gene: MIPS:YPR147C

A:Cross-references: SGD:S0006351

A:Map position: 16R

C:Keywords: transmembrane protein

F:160-176/Domain: transmembrane #status predicted <TMM>

Query Match 97.1%; Score 33; DB 2; Length 304;  
 Best Local Similarity 83.3%; Pred. No. 81;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
 |||||  
 DB 31 LLIWIP 36

## RESULT 9

■ kappa chain - human

C:Species: Homo sapiens (man)

C:Date: 06-Mar-1994 #sequence\_revision 26-May-1995 #text\_change 21-Jan-2000

C:Accession: S40312

R:Klein, R.; Jernichen, R.; Zachau, H.G.

Eur. J. Immunol. 23, 3248-3271, 1993

A:Title: Expressed human immunoglobulin chi genes and their hypermutation.

A:Reference number: S40312; MUID:94080891; PMID:8258341

A:Accession: S40312

A>Status: preliminary; translation not shown

A:Molecule type: mRNA

A:Residues: 1-126 <KLE>

A:Cross-references: EMBL:X72422; NID:q441312; PID:q441313

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:29-108/Domain: immunoglobulin homology <IMM>

Query Match 94.1%; Score 32; DB 2; Length 126;  
 Best Local Similarity 66.7%; Pred. No. 53;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
 |||||  
 DB 4 LLIWIP 9

## RESULT 10

■

hypothetical protein BH1479 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C:Accession: G83834

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: G83834

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-200 <STO>

A:Cross-references: GB:AP001512; GB:BA000004; NID:q10174030; PIDN:BA05198.1; GSPDB:G

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH1479

Query Match 94.1%; Score 32; DB 2; Length 200;  
 Best Local Similarity 83.3%; Pred. No. 83;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
 |||||  
 DB 193 LLIWIP 198

## RESULT 11

■

hypothetical protein C15B12.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 04-Mar-2000

C:Accession: T15500

R:Nhan, M.

submitted to the EMBL Data Library, March 1995

A:Description: The sequence of C. elegans cosmid C15B12.

A:Reference number: Z18362

A:Accession: T15500

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-269 <NHA>

A:Cross-references: EMBL:U23529; NID:q746592; PID:q746594; PIDN:AAC46577.1; CESP:C15B

A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:C15B12.2  
A:Introns: 3/3; 49/3; 105/2; 200/2; 245/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein C15B12.2

Query Match  
Best Local Similarity 94.1%; Score 32; DB 2; Length 269;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
| | | | |  
DB 76 LLIWIP 81

RESULT 12  
G82822  
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 [similarity] - Xylella fastidiosa  
C:Species: Xylella-fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 03-Jun-2002  
Accession: G82822  
Synopsis: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequences  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: G82822  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-501 <SIM>  
A:Cross-references: GB:AE00384; GB:AE003849; NID:g9105127; PIDN:AAF83128.1; GSPDB:GN001  
A:Experimental source: strain 965c  
R:Simmons, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Brlones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carriro, D.M.; Carreir, H  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincanl, A.P.; Ferreira, A.J.S.  
Submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Medeira, A.M.B.N.; Medeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeiri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF0317  
A:Superfamily: NADH dehydrogenase (ubiquinone) chain 4  
A:Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match  
Best Local Similarity 94.1%; Score 32; DB 2; Length 501;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
| | | | |  
DB 9 LLIWIP 14

RESULT 13  
T19862  
hypothetical protein C40C9.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
Accession: T19862  
R:Hemby, C.  
submitted to the EMBL Data Library, March 1996  
A:Accession: T19862  
A:Reference number: Z19188  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-542 <MIL>

A:Cross-references: EMBL:Z70266; PIDN:CAA94206.1; GSPDB:GN00028; CESP:C40C9.2  
A:Experimental source: clone C40C9  
C:Genetics:  
A:Gene: CESP:C40C9.2  
A:Map position: X  
A:Introns: 11/3; 69/3; 123/2; 173/3; 216/1; 248/1; 279/1; 314/1; 350/3; 430/1; 464/1;  
C:Superfamily: acetylcholine receptor

Query Match  
Best Local Similarity 94.1%; Score 32; DB 2; Length 542;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
| | | | |  
DB 355 LLIWIP 360

RESULT 14  
S00996  
Ig kappa chain precursor V region (A10) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 23-Jul-1999  
Accession: S00996  
R:Straubinger, B.; Thiebe, R.; Huber, C.; Osterholzer, E.; Zachau, H.G.  
Biol. Chem. Hoppe-Seyler 369, 601-607, 1988  
A:Title: Two unusual human immunoglobulin V-kappa genes.  
A:Reference number: S00996; MUID:89134397; PMID:2852016  
A:Accession: S00996  
A:Molecule type: DNA  
A:Residues: 1-114 <STR>  
A:Cross-references: EMBL:M27750; NID:q185914; PIDN:AAAS8912.1; PID:g553479  
A:Note: this sequence was determined from the germline gene  
C:Genetics:  
A:Introns: 16/1  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-114/Product: Ig kappa chain V region #status predicted <MAT>  
F:42-107/Disulfide bonds: #status predicted

Query Match  
Best Local Similarity 91.2%; Score 31; DB 2; Length 114;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
| | | | |  
DB 10 LLIWIP 15

RESULT 15  
S10146  
Ig kappa chain precursor V region (A14) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 21-Jan-2000  
Accession: S10146  
R:Straubinger, B.; Thiebe, R.; Huber, C.; Osterholzer, E.; Zachau, H.G.  
Biol. Chem. Hoppe-Seyler 369, 601-607, 1988  
A:Title: Two unusual human immunoglobulin V-kappa genes.  
A:Reference number: S00996; MUID:89134397; PMID:2852016  
A:Accession: S10146  
A:Molecule type: DNA  
A:Residues: 1-115 <STR>  
A:Cross-references: EMBL:M27751; NID:q185916; PIDN:AAAS8913.1; PID:g185917  
A:Note: this sequence was determined from the germline gene  
C:Genetics:  
A:Introns: 17/1  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-115/Product: Ig kappa chain V region #status predicted <MAT>  
F:36-110/Domain: immunoglobulin homology <IMM>  
F:43-108/Disulfide bonds: #status predicted

Query Match 91.28; Score 31; DB 2; Length 115;  
Best Local Similarity 66.7%; Pred. No. 74;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11:1:1  
Db 11 LLIWIP 16

Search completed: January 3, 2003, 15:33.15  
Job time : 11.9545 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:22:10 : Search time 5.18182 Seconds  
(without alignments)  
48.025 Million cell updates/sec

Title: US-09-543-188A-3  
Perfect score: 34  
Sequence: 1 LLIWIP 6

Scoring table: BIOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues  
al number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	402	1	YTCB_BACSU
2	34	100.0	419	1	O56A_DROME
3	34	100.0	2452	1	RPB1_PLARD
4	33	97.1	171	1	BB19_RABIT
5	31	91.2	131	1	KV3I_MOUSE
6	31	91.2	132	1	KV3F_MOUSE
7	31	91.2	335	1	MTFL_KLUTA
8	31	91.2	349	1	MRAY_CHLPP
9	31	91.2	365	1	MAP3_SCHPO
10	31	91.2	374	1	FCG1_HUMAN
11	31	91.2	394	1	EMRD_ECOLI
12	31	91.2	404	1	FCG1_MOUSE
13	31	91.2	443	1	DCVA_HELPY
14	31	91.2	443	1	DCVA_HELPY
15	30	88.2	115	1	KV3I_HUMAN
16	30	88.2	115	1	KV3I_MOUSE
17	30	88.2	116	1	KV3I_HUMAN
18	30	88.2	117	1	KV3I_HUMAN
19	30	88.2	117	1	YCC1_YEAST
20	30	88.2	128	1	KV3K_HUMAN
21	30	88.2	129	1	KV3K_HUMAN
22	30	88.2	129	1	KV3I_HUMAN
23	30	88.2	129	1	KV3M_HUMAN
24	30	88.2	164	1	Y40B_HISIN
25	30	88.2	240	1	P4SC_HAEIN
26	30	88.2	310	1	YGR1_YEAST
27	30	88.2	319	1	NUIM_POLOR
28	30	88.2	337	1	MPFL_SACKL
29	30	88.2	348	1	Y567_HELPY
30	30	88.2	349	1	Y567_HELPY
31	30	88.2	392	1	CEMA_NEPOL
32	30	88.2	467	1	GAC3_HUMAN
33	30	88.2	467	1	GAC3_MOUSE

34	30	88.2	467	1	GAC3_RAT	P28473	rattus norv
35	30	88.2	503	1	ALG6_CAEEL	O09226	caenorhabdi
36	30	88.2	575	1	ITR1_SCHPO	O10286	schizosacch
37	30	88.2	631	1	YCI0_ECOLI	P45848	eschrichia
38	30	88.2	710	1	IGAA_SALTY	P58721	salmonella
39	30	88.2	710	1	IGAA_SALTY	O9acp0	salmonella
40	30	88.2	711	1	IGAA_SALTY	P58720	eschrichia
41	30	88.2	711	1	IGAA_SALTY	P45800	eschrichia
42	30	88.2	1541	1	MRP2_RAT	O63120	rattus norv
43	30	88.2	1545	1	MRP2_HUMAN	O92887	homo sapien
44	30	88.2	1564	1	MRP2_RABIT	O28689	oryctolagus
45	29	85.3	115	1	MYHA_MOUSE	O61879	mus musculu

## ALIGNMENTS

RESULT 1  
YTCB\_BACSU STANDARD: PRT: 402 AA.  
ID P37482;  
AC 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical transport protein ytcB.  
GN ytcB.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=96051385; PubMed=7584024;  
RA Ogasawara N., Nakai S., Yoshikawa H.;  
RT "Systematic sequencing of the 180 kilobase region of the Bacillus  
subtilis chromosome containing the replication origin.";  
RL DNA Res. 1:1-14(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
Brouillet S., Brusch C.V., Caldwell B., Capano V., Carter N.M.,  
Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,  
Dentzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,  
Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
Guiseppi G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henaut A.,  
Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
Joris B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,  
Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,  
Kurita K., Laplus A., Lardinois S., Lauber J., Lazarevic V.,  
Lee S.M., Levine A., Liu H., Masuda S., Mauel J., Medigue C.,  
Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,  
Paro V., Pohl T.M., Portetlelle D., Porwollik S., Prescott A.M.,  
Priesen E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,  
Rieger M., Rivolta C., Roche E., Roche B., Rose M., Satale Y.,  
Sato T., Scanlan E., Schleich S., Schroeder R., Scoffone F.,  
Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
Takeuchi M., Yamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarelli A.,  
Viari A., Wambitt R., Wedler H., Wedler H., Weitzenecker T.,  
Winners P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
Yoshida K., Yoshikawa H.F., Zumslein E., Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
subtilis.";  
RL Nature 390:249-256(1997).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).

CC -1- SIMILARITY: BELONGS TO THE TC 2.A.1.17.1 TRANSPORTER FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: D26185; BAA05179.1; -  
DR EMBL: Z99124; CAB16085.1; -  
DR Subtilisin; Bg10007; yycB.  
DR InterPro: IPR004747; Cyan\_transport.  
DR TIGRfams: TIGR00896; Cynx; 1.  
KM Hypothetical protein; Transport: Transmembrane; Complete proteome.  
FT TRANSMEM 12 32  
FT TRANSMEM 48 68 POTENTIAL.  
FT TRANSMEM 80 100 POTENTIAL.  
FT TRANSMEM 101 121 POTENTIAL.  
FT TRANSMEM 134 154 POTENTIAL.  
FT TRANSMEM 168 188 POTENTIAL.  
FT TRANSMEM 212 232 POTENTIAL.  
FT TRANSMEM 248 268 POTENTIAL.  
FT TRANSMEM 291 311 POTENTIAL.  
FT TRANSMEM 339 359 POTENTIAL.  
FT TRANSMEM 367 387 POTENTIAL.  
SO SEQUENCE 402 AA; 43184 MW; 45C3348FE2183CD CRC64;  
  
Query Match 100.0%; Score 34; DB 1; Length 402;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 LTIWIP 6  
DB 181 LTIWIP 186  
  
RESULT 2  
056A\_DROME STANDARD: PRT: 419 AA.  
AC Q9V877;  
DT 16-OCT-2001 (Rel. 40; Created)  
DT 16-OCT-2001 (Rel. 40; Last sequence update)  
DT 15-JUN-2002 (Rel. 41; Last annotation update)  
DE Putative odorant receptor 56a.  
GN OR56A OR CG12501.  
CC Drosophila melanogaster (Fruit fly).  
CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
OC NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Cealinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H.C., Blazek R.G., Champs M., Pfeiffer B.D.,  
RA Wan K.H., Doyle G., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brodtler P.,  
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Dubin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Foster C., Gierl A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
RA Jaitli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RN Science 287:2185-2195(2000).  
RL [2]  
RP REVISIONS.  
RC STRAIN-Berkeley;  
RC Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Cealinker S.E.,  
RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,  
RA Harris N.L., Krommiller B., Marshall B., Millburn G.H., Richter J.,  
RA Russo S., Seale S.M.J., Smith E., Shu S., Smutnick F.,  
RA Whitfield E.J., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Mungall C.J., Lewis S.E.;  
RT "Annotation of Drosophila melanogaster genome.";  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: PROBABLE ROLE IN THE ODORANT RESPONSE, BEING AN ODORANT  
CC RECEPTOR.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -1- SIMILARITY: BELONGS TO FAMILY DR-OR OF G-PROTEIN COUPLED  
CC RECEPTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: AE003794; AAF57517.2; -  
DR Flybase: FBgn0034473; Or56a.  
DR InterPro: IPR004117; 7tm\_6.  
DR Pfam: PF02949; 7tm\_6; 1.  
KM Hypothetical protein; Transmembrane; G-protein coupled receptor;  
KW Olfaction; Multigene family.  
FT TRANSMEM 1 41  
FT TRANSMEM 42 62 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 63 76 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 77 97 2 (POTENTIAL).  
FT DOMAIN 98 137 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 138 158 3 (POTENTIAL).  
FT TRANSMEM 159 196 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 197 217 4 (POTENTIAL).  
FT DOMAIN 218 292 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 293 313 5 (POTENTIAL).  
FT DOMAIN 314 323 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 324 344 6 (POTENTIAL).  
FT DOMAIN 345 389 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 390 410 7 (POTENTIAL).  
FT DOMAIN 411 419 CYTOPLASMIC (POTENTIAL).  
SO SEQUENCE 419 AA; 48920 MW; F5F8254502560040 CRC64;  
  
Query Match 100.0%; Score 34; DB 1; Length 419;  
Best Local Similarity 100.0%; Pred. No. 27;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LTIWIP 6  
 |||||  
 Db 140 LTIWIP 145

RESULT 3  
 REP1\_PLAFD  
 ID REP1\_PLAFD STANDARD; PRT: 2452 AA.

AC P14248;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE DNA-directed RNA polymerase II largest subunit (EC 2.7.7.6).  
 GN RPII.  
 OS Plasmodium falciparum (isolate CDC / Honduras).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5836;  
 [1]

SEQUENCE FROM N.A.

RA MEDLINE-90098832; PubMed-2690004;  
 Li W.B., Beik D.J., Gu H., Tanaka M., Fox B.A., Inselburg J.;  
 RT "An enlarged largest subunit of Plasmodium falciparum RNA polymerase  
 RT II defines conserved and variable RNA polymerase domains.";  
 RL Nucleic Acids Res. 17:9621-9636(1989).  
 CC -1 FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 CC SUBSTRATES.

CC -1 CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA}(N).  
 CC -1 SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM TEN TO  
 CC TWELVE DIFFERENT POLYPEPTIDES. THIS POLYPEPTIDE IS THE LARGEST  
 CC COMPONENT OF RNA POLYMERASE II.  
 CC -1 SUBCELLULAR LOCATION: Nuclear.

CC -1 PTM: THE TANDAM 7 RESIDUES REPEATS CAN BE HIGHLY PHOSPHORYLATED.  
 CC THE PHOSPHORYLATION ACTIVATES POL2.  
 CC -1 MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES WERE  
 CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA  
 CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE  
 CC III FOR 5S AND TRNA GENES.

CC -1 SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

CC EMBL: X16561; CAA34560.1; -  
 DR PIR: S07485; RNZ02L.  
 DR InterPro: IPR000684; RNA\_polII\_repeat.  
 DR InterPro: IPR000722; RNA\_pol\_A.  
 DR InterPro: IPR002879; RNA\_pol\_A2.  
 DR Pfam: PF00623; RNA\_pol\_A: 1.  
 DR Pfam: PF01854; RNA\_pol\_A2: 1.  
 DR PROSITE: PS00115; RNA\_POL\_II\_REPEAT: 9.  
 KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;  
 KW DNA-binding; Nuclear protein; C2H2-type (POTENTIAL); Zinc-finger.  
 FT ZN\_FING 68 84  
 FT ZN\_FING 378 411  
 FT ZN\_FING 707 725  
 FT DOMAIN 1093 1128  
 FT DOMAIN 1144 1159  
 FT DOMAIN 1182 1193  
 FT DOMAIN 1687 1694  
 FT DOMAIN 1258 1290  
 FT DOMAIN 1261 1290  
 FT DOMAIN 1602 1612  
 FT DOMAIN 1746 1759  
 FT DOMAIN 1806 1820  
 FT DOMAIN 1806 1820

FT DOMAIN 2061 2246 HIGHLY DIVERGED HEPTAPEPTIDE REPEATS.  
 FT DOMAIN 2247 2384 CARBOXYL-TERMINAL 7-RESIDUE REPEATS.  
 SQ SEQUENCE 2452 AA; 278166 MW; F995E117F617A48F CRC64;

Query Match  
 Best Local Similarity 100.0%; Score 34; DB 1; Length 2452;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTIWIP 6  
 |||||  
 Db 582 LTIWIP 587

RESULT 4  
 BB19\_RABIT  
 ID BB19\_RABIT STANDARD; PRT: 171 AA.

AC Q05005;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Brush border 19.3 kDa protein precursor.  
 GN ADRAB-C.

OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE-93286138; PubMed-8509424;  
 RA Boli W., Schimid-Chanda T., Semenza G., Mantel N.;

RT "Messenger RNAs expressed in intestine of adult but not baby rabbits.  
 RT Isolation of cognate cDNAs and characterization of a novel brush  
 RT border protein with esterase and phospholipase activity.";  
 RL J. Biol. Chem. 268:12901-12911(1993).

CC -1 SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1 TISSUE SPECIFICITY: INTESTINE.  
 CC -1 DEVELOPMENTAL STAGE: EXPRESSED IN THE INTESTINE OF ADULT BUT NOT  
 CC BABY RABBITS.  
 -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

DR EMBL: Z12842; CAA78304.1; -  
 DR PIR: C45665; C45665.  
 KW Signal; Transmembrane.

FT SIGNAL 1 21  
 FT CHAIN 22 171  
 FT TRANSMEM 49 69  
 FT TRANSMEM 88 108  
 FT TRANSMEM 131 151  
 FT SEQUENCE 171 AA; 19290 MW; 1628238A011F9ADB CRC64;

Query Match  
 Best Local Similarity 97.18; Score 33; DB 1; Length 171;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LTIWIP 6  
 |||||  
 Db 63 LTIWIP 68

RESULT 5  
 KV31\_MOUSE  
 ID KV31\_MOUSE STANDARD; PRT: 131 AA.

AC P01661;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig kappa chain V-III region MOPC 63 precursor.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE OF 1-35.  
RX MEDLINE=78235887; PubMed=98179;  
RA Birstein Y., Schechter I.;  
RT "Primary structures of N-terminal extra peptide segments linked to  
RT the variable and constant regions of immunoglobulin light chain  
RT precursors: implications on the organization and controlled  
RT expression of immunoglobulin genes.";  
RL Biochemistry 17:2392-2400(1978).  
RN [2]  
RP SEQUENCE OF 21-131.  
RX MEDLINE=73140225; PubMed=4691517;  
RA McKean D.J., Potter M., Hood L.E.;  
RT "Mouse immunoglobulin chains. Pattern of sequence variation among  
RT kappa chains with limited sequence differences.";  
RL Biochemistry 12:760-771(1973).  
RN [3]  
RP REVISIONS.  
RX MEDLINE=79012520; PubMed=99744;  
RA McKean D.J., Bell M., Potter M.;  
RT "Mechanisms of antibody diversity: multiple genes encode structurally  
RT related mouse kappa variable regions.";  
RL Proc. Natl. Acad. Sci. U.S.A. 75:3913-3917(1978).  
DR HSP; P01679; 2FBU.  
DR PIR; A01935; KWSM6.  
DR DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig\_V.  
DR SMART: SM00406; IGV; 1.  
KM Immunoglobulin V region; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 131 IG KAPPA CHAIN V-III REGION MOPC 63.  
FT DOMAIN 21 43 FRAMEWORK-1.  
FT DOMAIN 44 58 COMPLEMENTARITY-DETERMINING-1.  
FT DOMAIN 59 73 FRAMEWORK-2.  
FT DOMAIN 74 80 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 81 112 FRAMEWORK-3.  
FT DOMAIN 113 121 COMPLEMENTARITY-DETERMINING-3.  
FT DOMAIN 122 131 FRAMEWORK-4.  
FT DISULFID 43 112 BY SIMILARITY.  
FT NON\_TER 131 131  
SQ SEQUENCE 131 AA: 14291 MW: D212EC9F08DC880A CRC64;  
Query Match 91.2%; Score 31; DB 1; Length 131;  
Best Local Similarity 66.7%; Pred. No. 33;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLIWIP 6  
DB 11 LLLWVP 16  
ID KVF\_MOUSE STANDARD; PRT; 132 AA.  
AC P01658;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig kappa chain V-III region MOPC 321 precursor.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE OF 1-37.  
RX MEDLINE=78235887; PubMed=98179;  
RA Birstein Y., Schechter I.;

RT "Primary structures of N-terminal extra peptide segments linked to  
RT the variable and constant regions of immunoglobulin light chain  
RT precursors: implications on the organization and controlled  
RT expression of immunoglobulin genes.";  
RL Biochemistry 17:2392-2400(1978).  
RN [2]  
RP SEQUENCE OF 21-132.  
RX MEDLINE=73140224; PubMed=4120629;  
RA McKean D.J., Potter M., Hood L.E.;  
RT "Mouse immunoglobulin chains. Partial amino acid sequence of a kappa  
RT chain.";  
RL Biochemistry 12:749-759(1973).  
CC -1- MISCELLANEOUS: THE PARTIAL SEQUENCE OF THE C REGION OF THIS  
CC BENCE-JONES PROTEIN WAS ALSO DETERMINED. IT DIFFERS FROM THAT  
CC REPORTED FOR MOUSE MOPC 21 ONLY IN THE TRANSPOSITION OF TWO NEARBY  
CC RESIDUES.  
DR PIR; A01933; KWS32.  
DR HSP; P01679; 2FBU.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig\_V.  
DR SMART: SM00406; IGV; 1.  
KM Immunoglobulin V region; Bence-Jones protein; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 132 IG KAPPA CHAIN V-III REGION MOPC 321.  
FT DOMAIN 21 43 FRAMEWORK-1.  
FT DOMAIN 44 58 COMPLEMENTARITY-DETERMINING-1.  
FT DOMAIN 59 73 FRAMEWORK-2.  
FT DOMAIN 74 80 FRAMEWORK-3.  
FT DOMAIN 81 112 COMPLEMENTARITY-DETERMINING-2.  
FT DOMAIN 113 121 COMPLEMENTARITY-DETERMINING-3.  
FT DOMAIN 122 131 FRAMEWORK-4.  
FT DISULFID 43 112 BY SIMILARITY.  
FT NON\_TER 132 132  
SQ SEQUENCE 132 AA: 14523 MW: 9F3B809BB773FEE9 CRC64;  
Query Match 91.2%; Score 31; DB 1; Length 132;  
Best Local Similarity 66.7%; Pred. No. 33;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLIWIP 6  
DB 11 LLLWVP 16  
ID MTF1\_KL0LA STANDARD; PRT; 335 AA.  
AC P87250;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Mitochondrial replication protein MTF1 (Mitochondrial transcription  
DE factor MTFB).  
GN MTF1.  
OS Kluyveromyces fragilis (Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.  
NCBI\_TaxID=28965;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97339478; PubMed=9196077;  
RA Carodenas J.A., Yun S., Shadel G.S., Clayton D.A., Bogenhagen D.F.;  
RT "Functional conservation of yeast mtrfb despite extensive sequence  
RT divergence.";  
RL Gene Expr. 6:219-230(1996).  
CC -1- FUNCTION: ESSENTIAL FOR MITOCHONDRIAL REPLICATION AND  
CC TRANSCRIPTION. CONFERS SELECTIVE PROMOTER RECOGNITION ON THE CORE  
CC SUBUNIT OF THE YEAST MITOCHONDRIAL RNA POLYMERASE. INTERACTS WITH  
CC DNA IN A NON-SPECIFIC MANNER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).  
CC -1- SIMILARITY: FAINT SIMILARITY WITH THE T4 GENE 32 PRODUCT AND WITH  
CC BACTERIAL SIGMA FACTORS.

```

CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; 081620; AAC49739.1; -
CC      Mitochondion; DNA replication; transcription regulation; DNA-binding.
CC      SW      SEQUENCE 335 AA; 39201 MW; 1B6EC6D7C6C3B791 CRC64;
SO
OY      Query Match      91.2%; Score 31; DB 1; Length 335;
OY      Best Local Similarity 66.7%; Pred. NO. 79;
OY      Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY      1 LIIWIP 6
OY      171 MLVWIP 176
OY      -----
RESULT 8
MRAY_CHLIPN
ID      MRAY_CHLIPN      STANDARD;      PRT;      349 AA.
AC      Q92706; Q9J512;
DT      30-MAY-2000 (Rel. 39, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Phospho-N-acetylmutamoyl-pentapeptide-transferase (EC 2.7.8.13) (UDP-
DE      MutNac-pentapeptide phosphotransferase).
GN      MRAY OR CPN0500 OR CP0966.
OC      Chlamydia pneumoniae (Chlamydia phyla pneumoniae).
OC      Bacteria; Chlamydiales; Chlamydiales; Chlamydiaceae; Chlamydia.
OX      NCBI_TaxID=833558;
OX      [1]
RN      SEQUENCE FROM N.A.
RP      STRAIN=CWL029;
RC      MEDLINE=99206606; PubMed=10192388;
RA      Kaitman S., Mitchell J., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA      Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT      "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL      Nat. Genet. 21:385-389(1999).
RN      [2]
RN      SEQUENCE FROM N.A.
RP      STRAIN=AR39;
RC      MEDLINE=20150255; PubMed=10684935;
RA      Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA      White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA      Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA      Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA      Eisen J., Fraser C.M.;
RT      "Genome sequences of Chlamydia trachomatis Mohn and Chlamydia
RT      pneumoniae AR39.";
RL      Nucleic Acids Res. 28:1397-1406(2000).
RN      [3]
RN      SEQUENCE FROM N.A.
RP      STRAIN=J138;
RC      MEDLINE=20330349; PubMed=10871362;
RA      Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA      Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT      "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT      from Japan and CWL029 from USA.";
RL      Nucleic Acids Res. 28:2311-2314(2000).
CC      -!- FUNCTION: FIRST STEP OF THE LIPID CYCLE REACTIONS IN THE
CC      BIOSYNTHESIS OF THE CELL WALL PEPTIDOGLYCAN (BY SIMILARITY).
CC      -!- CATALYTIC ACTIVITY: UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-L-
CC      lysyl-D-alanyl-D-alanine + undecaprenyl phosphate = UDP + N-
CC      acetylmuramoyl-L-alanyl-D-glutamyl-L-lysyl-D-alanyl-D-alanine-
CC      diposphoundecaprenol.
CC      -!- PATHWAY: Peptidoglycan biosynthesis.
CC      -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC      -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 4. MRAY

```

CC	SUBFAMILY.									
CC	-----									
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration									
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -									
CC	the European Bioinformatics Institute. There are no restrictions on its									
CC	use by non-profit institutions as long as its content is in no way									
CC	modified and this statement is not removed. Usage for and for commercial									
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/or_send_email_to_license@sib-sib.ch">http://www.isb-sib.ch/announce/or_send_email_to_license@sib-sib.ch</a> ).									
CC	-----									
DR	EMBL:	AE001670:	AAD19038.1:	-						
DR	EMBL:	AE002254:	AAF38746.1:	-						
DR	EMBL:	AP002548:	BAA39108.1:	-						
DR	TIGR:	CP0966:	-							
DR	InterPro:	IPR000715:	Glycos_transf_4.							
DR	InterPro:	IPR003524:	PNacCP_transf.							
DR	Pfam:	PF00953:	Glycos_transf_4: 1.							
DR	TIGRFAMS:	TIGR00445:	mrax: 1.							
DR	PROSITE:	PS01347:	MRAY_1: 1.							
DR	PROSITE:	PS01348:	MRAY_2: 1.							
KW	Peptidoglycan synthetis; Cell division; Transferase; Transmembrane;									
KW	Complete proteome.									
FT	TRANSMEM	13	33	POTENTIAL.						
FT	TRANSMEM	69	89	POTENTIAL.						
FT	TRANSMEM	91	111	POTENTIAL.						
FT	TRANSMEM	129	149	POTENTIAL.						
FT	TRANSMEM	165	185	POTENTIAL.						
FT	TRANSMEM	197	217	POTENTIAL.						
FT	TRANSMEM	228	248	POTENTIAL.						
FT	TRANSMEM	252	272	POTENTIAL.						
FT	TRANSMEM	278	298	POTENTIAL.						
FT	TRANSMEM	327	347	POTENTIAL.						
FT	CONFLICT	30	30	S -> A (IN REF. 1).						
SQ	SEQUENCE	349 AA:	38589 MM:	6CAA9283C594A88B CRC64:						
Query Match										
Best Local Similarity		91.2%:		Score 31:	DB 1:	Length 349:				
Matches		4: Conservative		2: Mismatches		0: Indels		0: Gaps		0:
OY	1	LIWIIP	6	: :						
DB	82	LIVWIP	87							
RESULT 9										
MAP3_SCHPO										
ID	MAP3_SCHPO	STANDARD:	PRT:	365	AA.					
AC	P31397:									
DT	01-JUL-1993	(rel. 26, Created)								
DT	01-JUL-1993	(rel. 26, Last sequence update)								
DT	15-JUN-2002	(rel. 41, Last annotation update)								
DE	Pheromone M-factor receptor.									
GN	MAP3 OR SPAC3P10.10C.									
OS	Schizosaccharomyces pombe (Fission yeast).									
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;									
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;									
OC	Schizosaccharomycetes.									
OX	NCBI_TaxID=4896;									
RN	[1]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN-ST711;									
RX	MEDLINE=93109361;	PubMed=8380233;								
RA	Tanaka K., Davey J., Imai Y., Yamamoto M.;									
RT	"Schizosaccharomyces pombe map3+ encodes the putative M-factor									
RL	receptor."									
RL	Mol. Cell. Biol.	13:80-88(1993).								
RN	[2]									
RP	SEQUENCE FROM N.A.									
RC	STRAIN=972;									
RX	MEDLINE=21848401;	PubMed=11859360;								
RA	Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,									
RA	Brooks K., Peat N., Hayles J., Baker S., Basham D., Bowman S.,									
RA	Stoughton K., Brown D., Brown S., Chillingworth T., Churcher C.M.,									

```

RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentiles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodson G.,
RA Holyoak S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., Melean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares K., Stevens K.,
RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
RA Weljens I., Vanstreels E., Rieger M., Scheafer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Egger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Motter S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Love T., McComble M.R., Paulsen I., Potashkin J.,
RA Spapakowski G.V., Ussery D., Bartell B.G., Nurse P.,
RA "The genome sequence of Schizosaccharomyces pombe." ;
RL Nature 415:871-880(2002).
CC -I- FUNCTION: RECEPTOR FOR THE PEPTIDE PHEROMONE M-FACTOR. A MATING
CC FACTOR OF S. POME. PHEROMONE SIGNALING IS ESSENTIAL FOR INITIATION
CC OF MEIOSIS IN S. POME. M-FACTOR SIGNALING ALONE MAY BE SUFFICIENT.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- INDUCTION: BY AT LEAST 3 TYPES OF REGULATION: THE MATING-TYPE OF
CC THE CELL, NUTRITIONAL CONDITIONS AND PHEROMONE SIGNALING.
CC -I- SIMILARITY: BELONGS TO FAMILY 4 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D10933; BAA01727.1; -.
DR EMBL: Z69369; CAA93308.1; -.
DR PIR: A48105; A48105.
DR InterPro: IPR001499; SRE3_GPCR.
DR Pfam: PF02076; SRE3.1.
DR Transmembrane: 7
KT FT TRANSMEM 24 POTENTIAL.
KT FT TRANSMEM 31 54 POTENTIAL.
KT FT TRANSMEM 73 100 POTENTIAL.
KT FT TRANSMEM 116 133 POTENTIAL.
KT FT TRANSMEM 155 182 POTENTIAL.
KT FT TRANSMEM 204 226 POTENTIAL.
KT FT TRANSMEM 265 283 POTENTIAL.
SQ SEQUENCE 365 AA; 42498 MW; 306A022FA7DCFEF7 CRC64;

Query Match 91.2%; Score 31; DB 1; Length 365;
Best Local Similarity 66.7%; Pred. NO. 85;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LIIWIP 6
|||:|
Db 294 LIIWIP 299

RESULT 10
FCG1_HUMAN
ID FCG1_HUMAN STANDARD; PRT: 374 AA.
AC P12314; P12315;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity immunoglobulin gamma fc receptor I precursor (Fc-gamma
RI) (FCRI) (IgG Fc receptor 1) (CD64).
GN FCGRIA OR FCGRI OR FCGI OR IGFI.

```

[illegible]

```

FT  VARSPLIC 333 374 HEKRVTSIQEDRLHEELKCOEQKEQLOEGVHKKEPOGA
FT  T -> GOALEAPTOGCA (IN ISOFORM B).
FT  VARIANT 26 26 L -> T
FT  VARIANT 59 59 /FTid=VAR_003953.
FT  VARIANT 59 59 N -> V.
FT  CONFLICT 25 25 /FTid=VAR_003954.
FT  SEQUENCE 374 AA; 42605 MW; 2C2AAB103ECF16E6 CRC64;
SQ
Query Match 91.2%; Score 31; DB 1; Length 374;
Best Local Similarity 66.7%; Pred. No. 87;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LLIWIP 6
Db 7 LLIWIP 12

RESULT 11
D_ECOLI STANDARD: PRT; 394 AA.
AC P31442;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein D.
GN EMDR OR B3673.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA MEDLINE=94059107; PubMed=8240355;
RA Nardolitskaya V., Schloesser M.J., Fan N.Y., Lewis R.;
RT "An E. coli gene emdr is involved in adaptation to low energy shock.";
RL Biochem. Biophys. Res. Commun. 196:803-809(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=K12 / MG1655;
RA MEDLINE=93315143; PubMed=7686882;
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
genome: organizational symmetry around the origin of replication.";
RL Genomics 16:551-561(1993).
CC -1- FUNCTION: MULTIDRUG RESISTANCE PUMP THAT PARTICIPATES IN A LOW
CC ENERGY SHOCK ADAPTATIVE RESPONSE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (potential).
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY). BELONGS TO THE
CC BCB/CMLA SUBFAMILY.
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L10328; AAA62025.1; ALT_INIT.
CC EMBL: AE000445; AAC76696.1; ALT_INIT.
CC Ecogene: EG11693; emrD.
CC InterPro: IPR004734; Drug_resist.
CC InterPro: IPR003652; Sub_transporter.
CC Pfam: PF00083; sugat.tr.1.
CC TIGRfams: TIGR00880; 2.A_01_02: 1.
CC Transport; Transmembrane; Inner membrane; Complete proteome.
CC TRANSMEM 9 29
CC TRANSMEM 47 67 POTENTIAL.
CC TRANSMEM 74 94 POTENTIAL.
CC TRANSMEM 96 116 POTENTIAL.

```

```

FT  TRANSMEM 135 155 POTENTIAL.
FT  TRANSMEM 163 183 POTENTIAL.
FT  TRANSMEM 213 233 POTENTIAL.
FT  TRANSMEM 243 263 POTENTIAL.
FT  TRANSMEM 277 297 POTENTIAL.
FT  TRANSMEM 299 319 POTENTIAL.
FT  TRANSMEM 330 350 POTENTIAL.
FT  TRANSMEM 365 385 POTENTIAL.
SQ
Query Match 91.2%; Score 31; DB 1; Length 394;
Best Local Similarity 83.3%; Pred. No. 91;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LLIWIP 6
Db 286 LLIWIP 291

RESULT 12
FCGI_MOUSE STANDARD: PRT; 404 AA.
ID FCGI_MOUSE
AC P26151;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE High affinity Immunoglobulin gamma FC receptor I precursor (FC-gamma
DE RI) (FCRI) (IgG FC receptor I).
GN FCGRI OR FCGI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90111035; PubMed=2136886;
RA Sears D.W., Osman N., Tate B., McKenzie I.F.C., Hogarth P.M.;
RT "Molecular cloning and expression of the mouse high affinity FC
RT receptor for IgG.";
RL J. Immunol. 144:371-378(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=92166399; PubMed=1531670;
RA Osman N., Kozak C.A., McKenzie I.F., Hogarth P.M.;
RT "Structure and mapping of the gene encoding mouse high affinity FC
RT gamma RI and chromosomal location of the human FC gamma RI gene.";
RL J. Immunol. 148:1570-1575(1992).
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN GAMMA. HIGH
CC AFFINITY RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: MACROPHAGE-SPECIFIC.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-10 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M31314; AAA40056.1; -.
CC PIR: A43511; A43511.
CC PIR: A46480; A46480.
CC HSSP: P12319; IALS.
CC MGP: MGI:95498; Fcgr1.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003598; Ig_c2.
CC InterPro: IPR003600; Ig_like.
CC Pfam: PF00047; Ig_3.
CC SMART: SM00410; Ig_like; 2.

```

```
DR SMART; SM00408; IGc2; 1.  
KW Igc-binding protein; Receptor; Transmembrane; Glycoprotein; Signal;  
KW Immunoglobulin domain; Repeat.  
FT SIGNAL 1 24  
FT CHAIN 25 404  
FT DOMAIN 25 297  
FT TRANSMEM 298 320  
FT DOMAIN 321 404  
FT DOMAIN 46 102  
FT DOMAIN 127 184  
FT DOMAIN 214 276  
FT DISULFID 53 95  
FT DISULFID 134 177  
FT DISULFID 221 269  
FT CARBOHYD 28 48  
FT CARBOHYD 69 69  
FT CARBOHYD 168 168  
FT CARBOHYD 249 249  
FT SEQUENCE 404 AA: 44887 MW: 1CAFF0038427677E7 CRC64;  
Query Match 91.2%; Score 31; DB 1; Length 404;  
Best Local Similarity 66.7%; Pred. No. 94;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 LLIWIP 6  
Db 16 LLWMP 21
```

## RESULT 13

```
DCUA_HELPJ STANDARD; PRT; 443 AA.  
ID DCUA_HELPJ  
AC 09ZLCO:
```

16-OCT-2001 (Rel. 40, Created)

16-OCT-2001 (Rel. 40, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)

DE Anaerobic C4-dicarboxylate transporter dcua.

DCUA OR JHP0660.

OS Helicobacter pylori J99 (Campylobacter pylori J99).

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

OC Helicobacter.

OX NCBI\_TaxID=85963;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=99120557; PubMed=9923682;  
Alm R.R., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Dolg P.C.,  
Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
Gibson R., Merberg D., Mills S.D., Jlang Q., Taylor D.E., Vovis G.F.,  
Trust T.J.;

RA "Genomic sequence comparison of two unrelated isolates of the human  
gastric pathogen Helicobacter pylori.";

RT Nature 397:176-180(1999).

RL -1- FUNCTION: RESPONSIBLE FOR THE TRANSPORT OF C4-DICARBOXYLATES FROM  
THE PERIPLASM ACROSS THE INNER MEMBRANE (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.

CC -1- SIMILARITY: BELONGS TO THE DCUA / DCUB (TC 2.A.13.1) FAMILY OF  
TRANSPORTERS.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AE001498; AAD06239.1; -  
DR InterPro: IPR004668; Dcu.  
DR Pfam: PF03605; Dcu; 1.  
DR TIGRFAMs: TIGR00770; Dcu; 1.

```
KW Transport; Transmembrane; Inner membrane; Complete proteome.  
FT TRANSMEM 1 20  
FT DOMAIN 21 21  
FT TRANSMEM 22 40  
FT TRANSMEM 41 56  
FT TRANSMEM 57 74  
FT TRANSMEM 75 88  
FT TRANSMEM 89 106  
FT TRANSMEM 107 135  
FT TRANSMEM 136 150  
FT TRANSMEM 151 221  
FT TRANSMEM 232 249  
FT TRANSMEM 250 267  
FT TRANSMEM 268 285  
FT TRANSMEM 286 295  
FT TRANSMEM 296 313  
FT TRANSMEM 314 335  
FT TRANSMEM 336 353  
FT TRANSMEM 354 358  
FT TRANSMEM 359 383  
FT TRANSMEM 384 443  
FT SEQUENCE 443 AA: 47291 MW: ECD39D53B250449 CRC64;  
Query Match 91.2%; Score 31; DB 1; Length 443;  
Best Local Similarity 66.7%; Pred. No. 1e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 LLIWIP 6  
Db 171 LLWMP 176
```

## RESULT 14

```
DCUA_HELPJ STANDARD; PRT; 443 AA.  
ID DCUA_HELPJ  
AC 025425:
```

15-DEC-1998 (Rel. 37, Created)

15-DEC-1998 (Rel. 37, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)

DE Anaerobic C4-dicarboxylate transporter dcua.

DCUA OR HP0724.

OS Helicobacter pylori (Campylobacter pylori).

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

OC Helicobacter.

OX NCBI\_TaxID=210;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=26695 / ATCC 700392;

RA MEDLINE=97394467; PubMed=9252185;  
Tomb J.-F., White O., Kervatage A.R., Clayton R.A., Sutton G.G.,  
Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,  
Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,  
Loftus B., Richardson D., Dodson R., Khakhria H.G., Glodek A.,  
McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.R.,  
Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,  
Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
Venter J.C.;

RA "The complete genome sequence of the gastric pathogen Helicobacter  
pylori.";

RT Nature 388:539-547(1997).

RL -1- FUNCTION: RESPONSIBLE FOR THE TRANSPORT OF C4-DICARBOXYLATES FROM  
THE PERIPLASM ACROSS THE INNER MEMBRANE (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.

CC -1- SIMILARITY: BELONGS TO THE DCUA / DCUB (TC 2.A.13.1) FAMILY OF  
TRANSPORTERS.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>



CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----  
CC EMBL: AE000585; AAD0773.1; -.  
DR TIGR: HP0724; -.  
DR InterPro: IPR004668; Dcu.  
DR Pfam: PF03605; Dcu; 1.  
DR TIGRFAMs: TIGR00770; Dcu; 1.  
KN Transport: Transmembrane; Inner membrane; Complete proteome.  
FT TRANSMEM 1 20  
FT DOMAIN 21 21  
FT TRANSMEM 22 21  
FT TRANSMEM 22 40  
FT TRANSMEM 41 56  
FT TRANSMEM 57 74  
FT TRANSMEM 75 88  
FT TRANSMEM 89 106  
FT TRANSMEM 107 135  
FT TRANSMEM 136 150  
FT TRANSMEM 151 231  
FT TRANSMEM 232 249  
FT TRANSMEM 250 267  
FT TRANSMEM 268 285  
FT TRANSMEM 286 295  
FT TRANSMEM 296 313  
FT TRANSMEM 314 335  
FT TRANSMEM 336 353  
FT TRANSMEM 354 358  
FT TRANSMEM 359 383  
FT TRANSMEM 384 443  
SQ SEQUENCE 443 AA; 47173 MW; 0E693630A0B50141 CRC64;  
  
Query Match  
Best Local Similarity 91.2%; Score 31; DB 1; Length 443;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLIWIP 6  
1:1:111  
Db 171 LMWIP 176  
  
RESULT 15  
KV31\_HUMAN STANDARD; PRT; 115 AA.  
AC P04433;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DI 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig kappa chain V-II region Vg precursor (Fragment).  
Homo sapiens (Human).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
OX [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=85087932; PubMed=6440122;  
RA Pech M., Zachau H.G.;  
RT "Immunoglobulin genes of different subgroups are interdigitated  
within the V<sub>K</sub> locus";  
RL Nucleic Acids Res. 12:9229-9236(1984).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: X01668; -; NOT\_ANNOTATED\_CDS.  
DR PIR: A01900; K3HUVG.  
DR HSSP: P80362; 1WTL.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003596; Ig\_V.  
DR Pfam: PF00047; Ig; 1.

DR SMART: SM00406; IGV; 1.  
KW Immunoglobulin V region; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 >115  
FT DOMAIN 21 43  
FT DOMAIN 44 54  
FT DOMAIN 55 69  
FT DOMAIN 70 76  
FT DOMAIN 77 108  
FT DOMAIN 109 115  
FT DISULFID 43 108  
FT NON\_TER 115 115  
SQ SEQUENCE 115 AA; 12575 MW; 2DE47CDA3A17D555 CRC64;  
  
Query Match  
Best Local Similarity 88.2%; Score 30; DB 1; Length 115;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLIWIP 6  
1:1:1  
Db 11 LLIWIP 16  
  
Search completed: January 3, 2003, 15:29:19  
Job time : 6.18182 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:24:35 ; Search time 20.3182 Seconds  
(without alignments)  
60.846 Million cell updates/sec

Title: US-09-543-188a-3  
Perfect score: 34  
Sequence: 1 LLIWIP 6

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_21:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mhvc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvtruss:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	157	5 Q9NFS8	Q9NFS8 Plasmodium
2	34	100.0	367	16 P71056	P71056 Bacillus su
3	34	100.0	554	4 Q92521	Q92521 homo sapien
4	34	100.0	554	4 Q8WVW7	Q8WVW7 homo sapien
5	34	100.0	754	10 Q92022	Q92022 arabidopsis
6	34	100.0	2457	5 Q77375	Q77375 plasmodium
7	33	97.1	133	16 Q985W5	Q985W5 rhizobium 1
8	33	97.1	290	17 Q87TC2	Q87TC2 methanosa
9	33	97.1	304	3 Q06522	Q06522 saccharomy
10	33	97.1	326	16 Q98K16	Q98K16 rhizobium 1
11	32	94.1	110	16 Q9KZB9	Q9KZB9 streptomy
12	32	94.1	138	2 Q936U1	Q936U1 pseudomon
13	32	94.1	139	2 Q05611	Q05611 pseudomon
14	32	94.1	200	16 Q9KC00	Q9KC00 bacillus ha
15	32	94.1	230	16 Q9RKX7	Q9RKX7 streptomy
16	32	94.1	269	5 Q18003	Q18003 caenorhabd

17	32	94.1	382	10 Q8W2P6	Q8W2P6 oryza sativ
18	32	94.1	479	2 Q8VTR2	Q8VTR2 bacteroides
19	32	94.1	501	16 Q9PG13	Q9PG13 xylella fas
20	32	94.1	542	5 Q18556	Q18556 caenorhabd
21	32	94.1	902	10 Q8S191	Q8S191 oryza sativ
22	32	94.1	3011	12 Q9DTP8	Q9DTP8 hepatitis c
23	31	91.2	100	4 Q9P0F3	Q9P0F3 homo sapien
24	31	91.2	103	4 Q9UMS9	Q9UMS9 homo sapien
25	31	91.2	131	16 Q9KXZ1	Q9KXZ1 bacillus ha
26	31	91.2	150	16 Q98LM8	Q98LM8 rhizobium 1
27	31	91.2	157	10 Q9C8R7	Q9C8R7 arabidopsis
28	31	91.2	159	11 Q8VD57	Q8VD57 mus musculu
29	31	91.2	186	4 Q9UMT0	Q9UMT0 homo sapien
30	31	91.2	229	5 Q19308	Q19308 caenorhabd
31	31	91.2	280	4 Q92637	Q92637 homo sapien
32	31	91.2	298	4 Q9HD97	Q9HD97 homo sapien
33	31	91.2	299	4 Q9UBK4	Q9UBK4 homo sapien
34	31	91.2	299	4 Q9UGN4	Q9UGN4 homo sapien
35	31	91.2	301	4 Q95100	Q95100 homo sapien
36	31	91.2	330	11 Q8R142	Q8R142 mus musculu
37	31	91.2	357	6 Q8SPW5	Q8SPW5 macaca fasc
38	31	91.2	374	4 Q92663	Q92663 homo sapien
39	31	91.2	375	4 Q92495	Q92495 homo sapien
40	31	91.2	396	16 Q8X541	Q8X541 escherichia
41	31	91.2	431	16 Q9K7K1	Q9K7K1 bacillus ha
42	31	91.2	494	5 Q9N983	Q9N983 leishmania
43	31	91.2	500	16 Q8P6N6	Q8P6N6 anabaena sp
44	31	91.2	521	5 Q18267	Q18267 caenorhabd
45	31	91.2	625	11 Q61468	Q61468 mus musculu

## ALIGNMENTS

RESULT 1  
Q9NFS8 PRELIMINARY: PRT: 157 AA.  
AC Q9NFS8:  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE RNA polymerase II (Fragment).  
GN RNAPOLII.  
OS Plasmodium yoelii.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5861;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=265BY;  
RX MEDLINE=20542027; PubMed=11087914;  
RA Nivez M.P., Achbarou A., Bienvenu J.D., Mazier D., Doerig C.,  
Vaquero C.;  
RT "A study of selected Plasmodium yoelii messengers RNAs during  
RT hepatocyte infection.";  
RL Mol. Biochem. Parasitol. 111:31-39(2000).  
DR EMBL, AJ271478; CAB70097.1; -;  
DR InterPro: IPR000722; RNA\_pol\_A.  
FT FTam: PF00623; RNA\_pol\_A: 1.  
FT NON\_TER 157  
SQ SEQUENCE 157 AA; 18667 MW; 79E0EA86CEAF5CA CRC64;

Query Match 100.0%; Score 34; DB 5; Length 157;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
|||||  
DB 66 LLIWIP 71

RESULT 2  
P71056

```
ID P71056 PRELIMINARY: PRT: 367 AA.
AC P71056: 008175;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein yveq.
GN yveq.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxId=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168TRP;
RA Fabret C., Quentin Y., Chapel N., Guisepi A., Halech J., Denizot F.;
RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RA Denizot F.C.;
RL Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.
[3]
RC STRAIN=168;
RA MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchart S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codan J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ertlan K.D., Ertlington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
RA Guisepi G., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Puje P., Furnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,
RA Saito T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terstera P., Tognoni A.,
RA Toseto V., Uchiyama S., Vandendol M., Vannier F., Vassarotti A.,
RA Viari A., Wambut R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
[4]
RN RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (Nov-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z71928; CA96474.1; -;
DR EMBL: Z94043; CAB08029.1; -;
DR EMBL: Z99121; CAB15436.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 367 AA; 42554 MW; 88D37986AFD9C9C6 CRC64;
```

```
Query Match 100.0%; Score 34; DB 16; Length 367;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 LTIWIP 6
DB 37 LTIWIP 42
```

```
RESULT 3
ID 092521 PRELIMINARY: PRT: 554 AA.
AC 092521;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)
DE Pig-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97015126; PubMed=8861954;
RA Takahashi M., Inoue N., Ohishi K., Maeda Y., Nakamura N., Endo Y.,
RA Fujita T., Takeda J., Kinoshita T.;
RT "Pig-B, a membrane protein of the endoplasmic reticulum with a large
RT lumenal domain, is involved in transferring the third mannose of the
RT GPI anchor."
RL EMO J. 15:4254-4261(1996).
DR EMBL: D42138; BAA07709.1; -;
SQ SEQUENCE 554 AA; 65056 MW; B2AF87D13ADFP90B3 CRC64;
```

```
Query Match 100.0%; Score 34; DB 4; Length 554;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 LTIWIP 6
DB 138 LTIWIP 143
```

```
RESULT 4
ID 08WVN7 PRELIMINARY: PRT: 554 AA.
AC 08WVN7;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Phosphatidylinositol glycan, class B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RA Straussberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC017711; AAH17711.1; -;
SQ SEQUENCE 554 AA; 64957 MW; E778418C02A24788 CRC64;
```

```
Query Match 100.0%; Score 34; DB 4; Length 554;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 LTIWIP 6
DB 138 LTIWIP 143
```

```
RESULT 5
ID 092022 PRELIMINARY: PRT: 754 AA.
AC 092022;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative LRR receptor-linked protein kinase (Putative LRR receptor-
DE like protein kinase).
```

```

GN F4C21.35 OR AT4G03390.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Habermann K., de la Bastide M., Huang E.N., Gnoc L., Schultz K.,
RA Preston R., Calma C., Martienssen R., Parnell L.D., Dechta N.,
RA McCombie W.R.;
RT "Arabidopsis thaliana BAC F4C21 from chromosome IV near 17 cM.";
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schultz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RT Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC005275; AAD14467.1; -.
DR EMBL; AL161496; CAB77824.1; -.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR002965; P_fich_extensn.
DR Pfam: PF00560; LRR; 5.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00370; LRR; 4.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR ATP-binding; Kinase; Receptor; Transferrase.
SQ SEQUENCE 754 AA; 82315 MW; 56B5946A597F63D CRC64;

Query Match
Best Local Similarity 100.0%; Score 34; DB 10; Length 754;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6
DB 19 LLIWIP 24

MULT 6
ID 077375 PRELIMINARY; PRT; 2457 AA.
AC 077375.
DT 01-NOV-1998 (TRENBLREL. 08, Created)
DT 01-NOV-1998 (TRENBLREL. 08, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE DNA-directed RNA polymerase II.
GN PFC0805W, MAL3P6.20.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RA MEDLINE=99376085; PubMed=10448855;
RA Bowman S., Lawson D., Basham D., Brown D., Chillingworth T.,
RA Churcher C.M., Craig A., Davies R.M., Devlin K., Feltwell T.,
RA Gentles S., Gilliam R., Hamlin N., Harris D., Holroyd S., Hornsby T.,
RA Horrocks P., Jagels K., Jassal B., Kyes S., McLean J., Moule S.,
RA Mungall K., Murphy L., Oliver K., Quail M.A., Rajandream M.-A.,
RA Ruter S., Skellton J., Squares R., Squares S., Sulston J.E.,
RA Whitehead S., Woodward J.R., Newbold C., Batteil B.G.;
RT "The complete nucleotide sequence of chromosome 3 of Plasmodium
falciparum .";

```

```

RL Nature 400:532-538(1999).
DR EMBL; Z98551; CAB1131.1; -.
DR InterPro: IPR000684; RNA_polII_repeat.
DR InterPro: IPR000722; RNA_pol_A.
DR InterPro: IPR002679; RNA_pol_A2.
DR InterPro: IPR002679; RNA_pol_A2.
DR Pfam: PF00623; RNA_pol_A; 1.
DR Pfam: PF01854; RNA_pol_A2; 1.
DR PROSITE; PS00115; RNA_POL_II_REPEAT; UNKNOWN_10.
SQ DNA-directed RNA polymerase.
QY SEQUENCE 2457 AA; 278674 MW; BB0FELL35F2FF43C CRC64;

Query Match
Best Local Similarity 100.0%; Score 34; DB 5; Length 2457;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6
DB 582 LLIWIP 587

RESULT 7
ID Q985W5 PRELIMINARY; PRT; 133 AA.
AC Q985W5.
DT 01-OCT-2001 (TRENBLREL. 18, Created)
DT 01-OCT-2001 (TRENBLREL. 18, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE Hypothetical protein mlr7498.
GN MLR7498.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFP303099;
RL MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto K., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti .";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003011; BAB53588.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 133 AA; 14309 MW; E2C038E7A39DDED CRC64;

Query Match
Best Local Similarity 97.1%; Score 33; DB 16; Length 133;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6
DB 93 LLIWIP 98

RESULT 8
ID Q8TTC2 PRELIMINARY; PRT; 290 AA.
AC Q8TTC2.
DT 01-JUN-2002 (TRENBLREL. 21, Created)
DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Predicted protein.
GN MA0515.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-C2A / ATCC 35395 / DSM 2834;  
RX MEDLINE-21929760; PubMed-11932238;  
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,  
RA Fitzhugh W., Calvo S., Engels R., Smirnov S., Atmoo D., Brown A.,  
RA Allen N., Naylor J., Stange-Thomann N., Deatellano K., Johnson R.,  
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,  
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,  
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,  
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reece J.N., Smith K.,  
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,  
RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,  
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
RA Metcalf W.M., Birren B.;  
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic  
and physiological diversity.";  
RT Genome Res. 12:532-542(2002).  
RL EMBL: AE010712; AM03959.1; -.  
DR EMBL: AE010712; AM03959.1; -.  
KW Complete proteome.  
SQ SEQUENCE 290 AA; 33697 MW; 931248B11C8E05CA CRC64;  
  
Query Match 97.1%; Score 33; DB 17; Length 290;  
Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 LLIWIP 6  
DB 189 LLIWIP 194  
  
RESULT 9  
O06522 PRELIMINARY; PRT; 304 AA.  
AC 006522;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Chromosome XVI COSMID 9659.  
GN YPR147C OR P9659.18.  
OS Saccharomyces cerevisiae (Baker's Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-97313271; PubMed-9169875;  
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W.,  
RA Araujo R., Aparicio A., Barrett B., Badcock K., Benes V., Botstein D.,  
RA Bowman S., Bruckner M., Carpenter J., Cherry J.M., Chung E.,  
RA Churcher C., Coster F., Davis K., Davis R.W., Dietrich F.S.,  
RA Drell M., Driaco T., Dubois E., Dusterhoft A., Duncan M., Floeth M.,  
RA Fortin N., Friesen J.D., Fritz C., Goffeau A., Hall J., Hebling U.,  
RA Heumann K., Hilbert H., Hillier L., Hunnicke-Smith S., Hyman R.,  
RA Johnson M., Kallman S., Kleine K., Komp C., Kudi O., Lashkari D.,  
RA Lew H., Lin A., Lin D., Louis E.J., Marathe R., Messenguy F.,  
RA Mees H.W., Milpach S., Moestl D., Muller-Auer S., Namath A.,  
RA Neutwich U., Oefner P., Pearson D., Petel F.X., Pohl T.M.,  
RA Punelle D., Schafer M., Scharfe M., Scherens B., Schramm S.,  
RA Schroeder M., Sidcu A.M., Tettein H., Urretazu L.A., Ushinsky S.,  
RA Viereckels F., Visers S., Voss H., Walsh S.V., Wambolt R., Wang Y.,  
RA Wedler E., Weller H., Winnelt E., Zhong W.W., Zollner A., Vo D.H.,  
RA Hani J.;  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";  
RL Nature 387:0-0(0).  
[2]  
RP SEQUENCE FROM N.A.  
RA Fulton L.;  
RN Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.  
[3]  
RP SEQUENCE FROM N.A.  
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,  
RA Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaha T.,  
RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,  
RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,

RA Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,  
RA Tach A., Trevasis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,  
RA Wilson R., Waterston R.; to the EMBL/Genbank/DBJ databases.  
RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.  
[4]  
RP SEQUENCE FROM N.A.  
RA Waterston R.;  
RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.  
[5]  
RP SEQUENCE FROM N.A.  
RA Jia Y., Cherry J.M.;  
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.  
DR EMBL: 040829; AAB68285.1; -.  
DR SGD: S0006351; YPR147C.  
DR InterPro: IPR000734; Lipase.  
DR InterPro: IPR000379; Ser\_estr-site.  
DR PROSITE: PS00120; LIPASE\_SER; UNKNOWN\_1.  
SQ SEQUENCE 304 AA; 34842 MW; B37EE95D46F44F34 CRC64;  
  
Query Match 97.1%; Score 33; DB 3; Length 304;  
Best Local Similarity 83.3%; Pred. No. 1.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 LLIWIP 6  
DB 31 LLIWIP 36  
  
RESULT 10  
O98K16 PRELIMINARY; PRT; 326 AA.  
AC 098K16;  
DT 01-OCT-2001 (TREMBLrel. 18, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Permease protein of ribose ABC transporter.  
GN MLI1679.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
[1]  
RP SEQUENCE FROM N.A.  
RX STRAIN-MAFE303099;  
RX MEDLINE-21082930; PubMed-11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
Mesorhizobium loti.";  
RL DNA Res. 7:331-338(2000).  
DR EMBL: AP002998; BAB48998.1; -.  
DR InterPro: IPR001851; Bac\_inmem\_transp.  
KW Complete proteome.  
SQ SEQUENCE 326 AA; 34344 MW; 934FF80EDAAAC8C4 CRC64;  
  
Query Match 97.1%; Score 33; DB 16; Length 326;  
Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 LLIWIP 6  
DB 176 LLIWIP 181  
  
RESULT 11  
O9K2H9 PRELIMINARY; PRT; 110 AA.  
AC 09K2H9;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

```

DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Putative integral membrane protein.
GN SC01166 OR SCG8A.20C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
CC Actinometales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bertley S.D., Chater K.F., Cerdano-Parraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Hatper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,
RA Hopwood D.A.;
*Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
Nature 417:141-147(2002).
DR EMBL, AL353663, CAB88950.1; -
SQ SEQUENCE 110 AA; 11809 MW; C3CCEFA7A44701297 CRC64;

Query Match          94.1%; Score 32; DB 16; Length 110;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6
Db 68 LLIWIP 73

RESULT 12
ID 093601 PRELIMINARY; PRT; 138 AA.
AC 093601;
DT 01-DEC-2001 (TRENBLREL. 19, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE Hypothetical 15.1 kDa protein.
OS Pseudomonas sp.
OG Plasmid pK14466.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LS46-6; TRANSPOSON=TN5046;
MEDLINE=21604134; PubMed=11763242;
RA Mindlin S.Z., Kholodil G.Y., Gorlenko Z.M., Minakhina S.V.,
RA Minakhin L.S., Kalyeva E.S., Kopteva A.V., Petrova M.A.,
RA Yurleva O.V., Nikiforov V.G.;
* Mercury resistance transposons of Gram-negative environmental
RT bacteria and their classification.";
RL Res. Microbiol. 152:811-822(2001).
DR EMBL, Y18360; CAC080080.1; -
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 138 AA; 15058 MW; 132E422A81681FE5 CRC64;

Query Match          94.1%; Score 32; DB 2; Length 138;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6
Db 112 LLIWIP 117

RESULT 13
ID 005611 PRELIMINARY; PRT; 139 AA.
AC 005611;

```

```

DT 01-JUL-1997 (TRENBLREL. 04, Created)
DT 01-JUL-1997 (TRENBLREL. 04, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE Transposon Tn5041 DNA.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KHP41; TRANSPOSON=TN5041;
MEDLINE=97419493; PubMed=9274008;
RA Kholodil G.Y., Yurleva O.V., Gorlenko Z.M., Mindlin S.Z., Bass I.A.,
RA Lomovskaya O.L., Kopteva A.V., Nikiforov V.G.;
* Tn5041 : a chimeric mercury resistance transposon closely related to
RT the toluene degradative transposon Tn4651.";
RL Microbiology 143:2549-2556(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KHP41; TRANSPOSON=TN5041;
RA Kholodil G.Y., Mindlin S.Z., Gorlenko Z.M., Bass I.A., Kalyeva E.S.,
RA Nikiforov V.;
* Host-dependent transposition of Tn5041.";
RL Russ. J. Genet. 36:365-373(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=KHP41; TRANSPOSON=TN5041;
RA Kholodil G.;
* Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL, X98999; CAA67452.1; -
SQ SEQUENCE 139 AA; 15114 MW; AF7F765C2C1FF5E3 CRC64;

Query Match          94.1%; Score 32; DB 2; Length 139;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6
Db 112 LLIWIP 117

RESULT 14
ID 09KC00 PRELIMINARY; PRT; 200 AA.
AC 09KC00;
DT 01-OCT-2000 (TRENBLREL. 15, Created)
DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE Hypothetical protein BH1479.
GN BH1479.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
CC Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
MEDLINE=20512582; PubMed=11058132;
RX Takami H., Nakasone K., Takaki Y., Ogasawara N., Kuhara S.,
RA Fuji F., Hirama C., Nakamura Y.,
RA Horikoshi K.;
* Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL, AP001512; BAB05198.1; -
DR InterPro; IPR003675; ABL.
DR Pfam; PF02517; ABL; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 200 AA; 22915 MW; F793D855F156C62D CRC64;

Query Match          94.1%; Score 32; DB 16; Length 200;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 LLIWIP 6  
 ||||:|  
 Db 193 LLIWIP 198

RESULT 15

O9RXX7 PRELIMINARY; PRT; 230 AA.  
 ID O9RXX7;  
 AC O9RXX7;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative integral membrane protein.  
 GN SC01418 OR SC6D7.21C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxId-1902;  
 RN [1]

SEQUENCE FROM N.A.  
 STRAIN-A3(2) / M145:  
 Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 Hopwood D.A.;  
 "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL133213; CAB61673.1;  
 SQ SEQUENCE 230 AA; 23916 MW; 5E87093F7CC1CC26 CRC64;

Query Match 94.1%; Score 32; DB 16; Length 230;  
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
 ||||:|  
 Db 78 LLIWIP 83

Search completed: January 3, 2003, 15:31:56  
 Job time : 22.3182 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:16 ; Search time 9 Seconds  
(without alignments)  
19.615 Million cell updates/sec

Title: US-09-543-188a-3

Perfect score: 34

Sequence: 1 LLIWIP 6

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PCtUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfill1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	91.2	20	3	US-08-996-139-16
2	31	91.2	20	4	US-08-995-659-16
3	31	91.2	20	4	US-09-215-649A-16
4	31	91.2	20	4	US-09-320-424-25
5	31	91.2	20	4	US-09-577-780-16
6	31	91.2	21	3	US-08-518-835-5
7	31	91.2	126	1	US-07-634-278-21
8	31	91.2	126	1	US-08-477-728-21
9	31	91.2	126	1	US-08-474-040-21
10	31	91.2	126	1	US-08-487-200-21
11	31	91.2	126	2	US-08-656-586-6
12	31	91.2	126	4	US-08-484-537-21
13	31	91.2	127	3	US-08-649-100-33
14	31	91.2	128	4	US-08-348-548-6
15	31	91.2	128	4	US-09-450-520A-6
16	31	91.2	128	5	PCT-US95-15716-6
17	31	91.2	131	1	US-07-634-278-67
18	31	91.2	131	1	US-08-477-728-67
19	31	91.2	131	1	US-08-474-040-67
20	31	91.2	131	1	US-08-487-200-67
21	31	91.2	131	1	US-08-137-117D-25
22	31	91.2	131	1	US-08-137-117D-33
23	31	91.2	131	2	US-08-436-717-25
24	31	91.2	131	2	US-08-436-717-33
25	31	91.2	131	2	US-08-621-751A-10
26	31	91.2	131	2	US-08-621-751A-14
27	31	91.2	131	3	US-08-589-939-3

28	31	91.2	131	3	US-08-836-561-25	Sequence 25, Appl
29	31	91.2	131	4	US-08-484-537-67	Sequence 16, Appl
30	31	91.2	131	4	US-08-579-378A-14	Sequence 67, Appl
31	31	91.2	131	4	US-08-579-378A-18	Sequence 18, Appl
32	31	91.2	132	2	US-08-379-057-16	Sequence 16, Appl
33	31	91.2	132	2	US-08-483-636-2	Sequence 2, Appl
34	31	91.2	133	5	PCT-US93-11611-9	Sequence 9, Appl
35	31	91.2	146	2	US-08-653-402B-12	Sequence 12, Appl
36	31	91.2	219	2	US-08-902-516-2	Sequence 2, Appl
37	31	91.2	240	4	US-09-301-593-36	Sequence 36, Appl
38	31	91.2	270	4	US-09-532-856-7	Sequence 7, Appl
39	31	91.2	354	4	US-09-393-627B-28	Sequence 28, Appl
40	31	91.2	19	1	US-07-977-696C-66	Sequence 66, Appl
41	30	88.2	19	1	US-08-129-930B-66	Sequence 66, Appl
42	30	88.2	19	4	US-08-976-288A-66	Sequence 66, Appl
43	30	88.2	24	1	US-08-162-102C-27	Sequence 27, Appl
44	30	88.2	24	1	US-08-162-102C-27	Sequence 47, Appl
45	30	88.2	24	1	US-08-162-102C-27	

#### ALIGNMENTS

RESULT 1  
US-08-996-139-16

; Sequence 16, Application US/08996139

; Patent No. 6017729

; GENERAL INFORMATION:

; APPLICANT: Anderson, Dirk M.

; APPLICANT: Galibert, Laurent

; APPLICANT: Maraskovsky, Eugene

; TITLE OF INVENTION: Receptor Activator of NF-kappaB

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESS: Immunex Corporation, Law Department

; STREET: 51 University Street

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98101

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: Apple Power Macintosh

; OPERATING SYSTEM: Apple Operating System 7.5.5

; SOFTWARE: Microsoft Word for Power Macintosh 6.0.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/996,139

; FILING DATE: 22 DECEMBER 1997

; CLASSIFICATION:

; APPLICATION NUMBER: US/08/996,139

; FILING DATE: 14 OCTOBER 1997

; APPLICATION NUMBER: US/08/813,509

; FILING DATE: 07 MARCH 1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: USSN 08/772,330

; FILING DATE: 23 DECEMBER 1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Perkins, Patricia Anne

; REGISTRATION NUMBER: 34,693

; REFERENCE/DOCKET NUMBER: 2851-A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206)567-0430

; TELEFAX: (206)233-0644

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 20 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-996-139-16

Query Match 91.2% Score 31; DB 3; Length 20;  
Best Local Similarity 66.7%; Pred. No. 21;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11 LLLWVP 16

## RESULT 2

US-08-995-659-16  
; Sequence 16, Application US/08995659  
; Patent No. 6242213  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Dirk M.  
; APPLICANT: Galibert, Laurent  
; APPLICANT: Maraskovsky, Eugene  
; TITLE OF INVENTION: Ligand for Receptor Activator of NF-kappaB  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation, Law Department  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: Apple Power Macintosh  
; SOFTWARE: Microsoft Word for Power Macintosh 6.0.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/995,659  
; FILING DATE: 22 DECEMBER 1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USSN 60/064,671  
; FILING DATE: 14 OCTOBER 1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USSN 08/813,509  
; FILING DATE: 07 MARCH 1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USSN 08/772,330  
; FILING DATE: 23 DECEMBER 1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia Anne  
; REGISTRATION NUMBER: 34,693  
; REFERENCE/DOCKET NUMBER: 2852-A  
; TELEPHONE: (206)587-0430  
; TELEFAX: (206)233-0644  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-995-659-16

Query Match 91.2% Score 31; DB 4; Length 20;  
Best Local Similarity 66.7%; Pred. No. 21;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11 LLLWVP 16

RESULT 3  
US-09-215-649A-16

; Sequence 16, Application US/09215649A  
; Patent No. 6271349  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Dirk M.  
; APPLICANT: Galibert, Laurent  
; APPLICANT: Maraskovsky, Eugene  
; TITLE OF INVENTION: Receptor Activator of NF-kappaB  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation, Law Department  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: Apple Power Macintosh  
; SOFTWARE: Microsoft Word for Power Macintosh 6.0.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/215,649A  
; FILING DATE: 17-Dec-1998  
; CLASSIFICATION: <unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/996,139  
; FILING DATE: <unknown>  
; APPLICATION NUMBER: USSN 08/813,509  
; FILING DATE: 07 MARCH 1997  
; APPLICATION NUMBER: USSN 08/772,330  
; FILING DATE: 23 DECEMBER 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Perkins, Patricia Anne  
; REGISTRATION NUMBER: 34,693  
; REFERENCE/DOCKET NUMBER: 2851-A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206)587-0430  
; TELEFAX: (206)233-0644  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-09-215-649A-16

Query Match 91.2% Score 31; DB 4; Length 20;  
Best Local Similarity 66.7%; Pred. No. 21;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11 LLLWVP 16

RESULT 4  
US-09-320-424-25  
; Sequence 25, Application US/09320424  
; Patent No. 6284236  
; GENERAL INFORMATION:  
; APPLICANT: Wiley, Steven R.  
; APPLICANT: Goodwin, Raymond G.  
; TITLE OF INVENTION: Cytokine that Induces Apoptosis  
; FILE REFERENCE: 2835-E  
; CURRENT APPLICATION NUMBER: US/09/320,424  
; EARLIER FILING DATE: 1999-05-26  
; EARLIER APPLICATION NUMBER: 09/190,046  
; EARLIER FILING DATE: 1998-11-10  
; EARLIER APPLICATION NUMBER: 09/048,641  
; EARLIER FILING DATE: 1998-03-26  
; EARLIER APPLICATION NUMBER: 08/670,354  
; EARLIER FILING DATE: 1996-06-25

EARLIER APPLICATION NUMBER: 08/548,368  
EARLIER FILING DATE: 1995-11-01  
EARLIER APPLICATION NUMBER: 08/496,632  
EARLIER FILING DATE: 1995-06-29  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 25  
LENGTH: 20  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-09-320-424-25

Query Match  
Best Local Similarity 91.2%; Score 31; DB 4; Length 30;  
66.7%; Pred. No. 21;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11:1:1  
DB 11 LLIWIP 16

SUPT 5  
US-09-577-780-16  
Sequence 16, Application US/09577780  
Patent No. 6419929  
GENERAL INFORMATION:  
APPLICANT: Anderson, Dirk M.  
Galibert, Laurent  
Maraskovsky, Eugene  
TITLE OF INVENTION: Ligand for Receptor Activator of NF-kappaB  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation, Law Department  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/577,780  
FILING DATE: 24-May-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/995,659  
FILING DATE: <Unknown>  
APPLICATION NUMBER: USSN 08/813,509  
FILING DATE: 07 MARCH 1997  
APPLICATION NUMBER: USSN 08/772,330  
FILING DATE: 23 DECEMBER 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,693  
REFERENCE/DOCKET NUMBER: 2852-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)233-0644  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-09-577-780-16

Query Match  
Best Local Similarity 91.2%; Score 31; DB 4; Length 20;  
66.7%; Pred. No. 21;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11:1:1  
DB 11 LLIWIP 16

RESULT 6  
US-08-518-835-5  
Sequence 5, Application US/08518835  
Patent No. 6017754  
GENERAL INFORMATION:  
APPLICANT: CHESTNUT, JONATHAN D.  
APPLICANT: HOEFLE, JAMES P.  
TITLE OF INVENTION: NOVEL SYSTEM FOR ISOLATING AND  
IDENTIFYING EUKARYOTIC CELLS TRANSFECTED WITH GENES AND  
TITLE OF INVENTION: VECTORS, HOST CELLS AND METHODS THEREOF  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 GARDEN CITY PLAZA  
CITY: GARDEN CITY  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/518,835  
FILING DATE: 24-AUG-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: DIGILIO, FRANK S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 11093  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-518-835-5

Query Match  
Best Local Similarity 91.2%; Score 31; DB 3; Length 21;  
66.7%; Pred. No. 22;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11:1:1  
DB 11 LLIWIP 16

RESULT 7  
US-07-634-278-21  
Sequence 21, Application US/07634278  
Patent No. 5530101  
GENERAL INFORMATION:  
APPLICANT: QUEEN, Cary L.  
APPLICANT: CO, Man Sung  
APPLICANT: SCHNEIDER, William P.  
APPLICANT: LANDOLFI, Nicholas F.  
APPLICANT: COELING, Kathleen L.  
APPLICANT: SELICK, Harold E.  
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourile and Crew  
STREET: 379 Lytton Avenue

CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US 07/634,278  
FILING DATE: 19-DEC-1990  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/590,274  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-002600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 126 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-634-278-21

Query Match 91.2%; Score 31; DB 1; Length 126;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
111111  
DB 11 LLIWIP 16

RESULT 8  
US-08-477-728-21  
Sequence 21, Application US/08477728  
Patent No. 5585089  
GENERAL INFORMATION:  
APPLICANT: QUEEN, Cary L.  
APPLICANT: SCHNEIDER, William P.  
APPLICANT: SELICK, Harold E.  
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,728  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/634,278  
FILING DATE: 19-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/590,274  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-002600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 126 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-477-728-21

Query Match 91.2%; Score 31; DB 1; Length 126;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
111111  
DB 11 LLIWIP 16

RESULT 9  
US-08-474-040-21  
Sequence 21, Application US/08474040  
Patent No. 5693761  
GENERAL INFORMATION:  
APPLICANT: QUEEN, Cary L.  
APPLICANT: CO, Man Sung  
APPLICANT: SCHNEIDER, William P.  
APPLICANT: LANDOLFI, Nicholas P.  
APPLICANT: COELINGH, Kathleen L.  
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,040  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/634,278  
FILING DATE: 19-DEC-1990  
APPLICATION NUMBER: US 07/590,274  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-002600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 126 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-474-040-21

Query Match 91.2%; Score 31; DB 1; Length 126;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11:1:1  
Db 11 LLLWVP 16

RESULT 10  
US-08-487-200-21  
Sequence 21, Application US/08487200  
Patent No. 5693762  
GENERAL INFORMATION:  
APPLICANT: QUEEN, Cary L.  
APPLICANT: CO, Man Sung  
APPLICANT: SCHNEIDER, William P.  
APPLICANT: LANDOLFI, Nicholas F.  
APPLICANT: COELINGH, Kathleen L.  
APPLICANT: SELICK, Harold E.  
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,200  
FILING DATE: 7-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/634,278  
FILING DATE: 19-DEC-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/590,274  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-002610  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400

TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 126 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-487-200-21

Query Match 91.2%; Score 31; DB 1; Length 126;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11:1:1  
Db 11 LLLWVP 16

RESULT 11  
US-08-656-586-6  
Sequence 6, Application US/08656586  
Patent No. 5834597  
GENERAL INFORMATION:  
APPLICANT: TSO, J. Yun  
APPLICANT: Cole, Michael S.  
APPLICANT: Anasetti, Claudio  
TITLE OF INVENTION: Mutated No. 5834597activating IgG2 Domains and  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/656,586  
FILING DATE: 31-MAY-1996  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebeschuetz, Joseph O.  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 11823-007210US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 126 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-656-586-6

Query Match 91.2%; Score 31; DB 2; Length 126;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11:1:1  
Db 11 LLLWVP 16

RESULT 12  
US-08-484-537-21  
Sequence 21, Application US/08484537  
Patent No. 6180370

GENERAL INFORMATION:  
APPLICANT: QUEEN, Cary L.  
APPLICANT: CO, Man Sung  
APPLICANT: SCHNEIDER, William P.  
APPLICANT: LANDOLFI, Nicholas F.  
APPLICANT: COELINGH, Kathleen L.  
APPLICANT: SELICK, Harold E.  
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourile and Crew  
STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,537  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/634,278  
FILING DATE: 19-DEC-1990  
APPLICATION NUMBER: US/07/590,274  
FILING DATE: 28-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/310,252  
FILING DATE: 13-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/290,975  
FILING DATE: 28-DEC-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 11823-002600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 126 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
08-484-537-21

Query Match 91.2%; Score 31; DB 4; Length 126;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
Db 11 LLIWIP 16

RESULT 13  
US-08-649-100-33  
Sequence 33, Application US/08649100  
Patent No. 6114507  
GENERAL INFORMATION:  
APPLICANT: SHIRAKAMA, KAMON  
APPLICANT: MATSUUE, TOMOKAZU  
APPLICANT: NAGATA, SHIGEKAZU  
APPLICANT: CO, MAN SUNG  
APPLICANT: VASQUEZ, MAXIMILIANO  
TITLE OF INVENTION: ANTI-FAS LIGAND ANTIBODY AND ASSAY  
TITLE OF INVENTION: METHOD USING THE ANTI-FAS LIGAND ANTIBODY  
NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH  
STREET: PO BOX 747  
CITY: FALLS CHURCH  
STATE: VA  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/649,100  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MORPHY JR, GERALD M  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1110-160  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 127 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-649-100-33

Query Match 91.2%; Score 31; DB 3; Length 127;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
Db 11 LLIWIP 16

RESULT 14  
US-08-348-548-6  
Sequence 6, Application US/08348548  
Patent No. 6258529  
GENERAL INFORMATION:  
APPLICANT: Berdoz, Jose  
APPLICANT: Kraehenbuhl, Jean Pierre  
TITLE OF INVENTION: PCR AMPLIFICATION OF REARRANGED GENOMIC  
TITLE OF INVENTION: VARIABLE REGIONS OF IMMUNOGLOBULIN GENES  
NUMBER OF SEQUENCES: 108  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street, Suite 3100  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/348,548  
FILING DATE: 01-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 06132/009001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-5070

TELEX: 200154  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 128 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-348-548-6

Query Match 91.2%; Score 31; DB 4; Length 128;  
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
 11:1:1  
 Db 11 LLLWVP 16

SU/T 15  
 09-450-520A-6  
 Sequence 6, Application US/09450520A  
 Patent No. 6329511  
 GENERAL INFORMATION:  
 APPLICANT: Vasquez, Maximiliano  
 APPLICANT: Landolfi, Nicholas F.  
 APPLICANT: Tsunashita, Naoya  
 APPLICANT: Queen, Cary L.  
 APPLICANT: Protein Design Labs, Inc.  
 TITLE OF INVENTION: Humanized Antibodies To Gamma-Interferon  
 FILE REFERENCE: 011823-008110US  
 CURRENT APPLICATION NUMBER: US/09/450,520A  
 PRIOR FILING DATE: 1998-11-29  
 PRIOR APPLICATION NUMBER: 60/110,523  
 PRIOR FILING DATE: 1998-12-01  
 NUMBER OF SEQ ID NOS: 13  
 SOFTWARE: Patent In Ver. 2.1  
 SEQ ID NO 6  
 LENGTH: 128  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: human-mouse  
 OTHER INFORMATION: transgenic construct HuZAP VL  
 US-09-450-520A-6

Query Match 91.2%; Score 31; DB 4; Length 128;  
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
 11:1:1  
 Db 11 LLLWVP 16

Search completed: January 3, 2003, 15:34:27  
 Job time : 10 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:32:11 : Search time 4.77273 Seconds  
(without alignments)  
23.825 Million cell updates/sec

Title: US-09-543-188a-3  
Perfect score: 34  
Sequence: 1 LITWIP 6

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 117078 seqs, 18951520 residues  
1 number of hits satisfying chosen parameters: 117078

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCR\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/PCRUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	100.0	87	10	US-09-864-761-40310
2	32	94.1	177	10	US-09-764-864-942
3	31	91.2	20	9	US-09-877-650-16
4	31	91.2	20	10	US-09-871-856-16
5	31	91.2	80	10	US-09-894-018-97
6	31	91.2	98	10	US-09-894-018-101
7	31	91.2	106	10	US-09-894-018-95
8	31	91.2	107	10	US-09-894-018-93
9	31	91.2	107	10	US-09-894-018-103
10	31	91.2	123	10	US-09-894-018-109
11	31	91.2	128	10	US-09-881-823-14
12	31	91.2	128	10	US-09-992-524-6
13	31	91.2	130	10	US-09-894-018-99
14	31	91.2	131	10	US-09-286-240-2
15	31	91.2	132	9	US-09-879-461-2
16	31	91.2	134	10	US-09-881-823-2
17	31	91.2	144	10	US-09-894-018-129
18	31	91.2	147	10	US-09-894-018-131
19	31	91.2	148	10	US-09-894-018-127

20	31	91.2	157	10	US-09-894-018-117	Sequence 117, App
21	31	91.2	168	10	US-09-894-018-115	Sequence 115, App
22	31	91.2	169	10	US-09-894-018-119	Sequence 119, App
23	31	91.2	180	10	US-09-894-018-125	Sequence 125, App
24	31	91.2	206	10	US-09-894-018-111	Sequence 111, App
25	31	91.2	207	9	US-10-077-438-3	Sequence 3, Appl1
26	31	91.2	207	9	US-10-077-137-3	Sequence 3, Appl1
27	31	91.2	207	9	US-10-080-797-3	Sequence 3, Appl1
28	31	91.2	211	10	US-09-894-018-85	Sequence 85, Appl
29	31	91.2	219	10	US-09-847-185-2	Sequence 2, Appl1
30	31	91.2	219	10	US-09-894-018-113	Sequence 113, App
31	31	91.2	238	9	US-09-903-327A-4	Sequence 4, Appl1
32	31	91.2	255	10	US-09-894-018-89	Sequence 89, Appl
33	31	91.2	276	10	US-09-894-018-141	Sequence 141, App
34	31	91.2	277	10	US-09-894-018-81	Sequence 81, Appl
35	31	91.2	280	10	US-09-894-018-83	Sequence 83, Appl
36	31	91.2	281	10	US-09-894-018-91	Sequence 91, Appl
37	31	91.2	288	9	US-09-895-913A-42	Sequence 42, Appl
38	31	91.2	308	10	US-09-894-018-105	Sequence 105, App
39	31	91.2	308	10	US-09-894-018-107	Sequence 107, App
40	31	91.2	399	10	US-09-925-101-1128	Sequence 1128, Ap
41	31	91.2	419	10	US-09-894-018-123	Sequence 123, App
42	31	91.2	446	10	US-09-792-2008-10	Sequence 10, App
43	31	91.2	446	10	US-09-792-2008-12	Sequence 12, Appl
44	31	91.2	456	10	US-09-894-018-121	Sequence 121, App
45	31	91.2	463	10	US-09-835-147-30	Sequence 30, Appl

## ALIGNMENTS

RESULT 1  
US-09-864-761-40310  
Sequence 40310, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
FILE REFERENCE: Aecmica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263,6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 40310  
LENGTH: 87  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL136308.4  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3  
OTHER INFORMATION: EST\_HUMAN HIT: AW085507.1, EVALUATE 6.40e+00  
OTHER INFORMATION: SWISSPROT HIT: Q80946, EVALUATE 2.60e+00  
US-09-864-761-40310

Query Match 100.0%; Score 34; DB 10; Length 87;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
|||||  
Db 48 LLIWIP 53

RESULT 2  
US-09-764-864-942  
Sequence 942; Application US/09764864  
Patent No. US20020132753A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT23  
CURRENT APPLICATION NUMBER: US/09/764,864  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 1792  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 942  
LENGTH: 177  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-864-942

Query Match 94.1%; Score 32; DB 10; Length 177;  
Best Local Similarity 83.3%; Pred. No. 63;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
|||||  
Db 35 LLIWIP 40

RESULT 3  
US-09-877-650-16  
Sequence 16; Application US/09877650  
Patent No. US20020169117A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Dirk M.  
Galibert, Laurent  
Maraskovsky, Eugene  
TITLE OF INVENTION: Ligand for Receptor Activator of NF-kappa

NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation, Law Department  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/877,650  
FILING DATE: 08-Jun-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/995,659  
FILING DATE: 1997-12-22  
APPLICATION NUMBER: USSN 08/813,509  
FILING DATE: 07 MARCH 1997  
APPLICATION NUMBER: USSN 08/772,330  
FILING DATE: 23 DECEMBER 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,693  
REFERENCE/DOCKET NUMBER: 2852-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)233-0644  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-09-877-650-16

Query Match 91.2%; Score 31; DB 9; Length 20;  
Best Local Similarity 66.7%; Pred. No. 13;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
|||:-|  
Db 11 LLIWIP 16

RESULT 4  
US-09-871-856-16  
Sequence 16; Application US/09871856  
Patent No. US20020081720A1  
GENERAL INFORMATION:  
APPLICANT: Anderson, Dirk M.  
Galibert, Laurent  
Maraskovsky, Eugene  
TITLE OF INVENTION: Receptor Activator of NF-kappa  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation, Law Department  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/871,856

```

; FILING DATE: 31-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/996,139
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US98 08/813,509
; FILING DATE: 07 MARCH 1997
; APPLICATION NUMBER: US98 08/772,330
; FILING DATE: 23 DECEMBER 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2851-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-871-856-16

Query Match          91.2%; Score 31; DB 10; Length 20;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6
|:|:|
DB 11 LLIWVP 16

RESULT 5
US-09-894-018-97
; Sequence 97, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV.3s2(-3)
US-09-894-018-97

Query Match          91.2%; Score 31; DB 10; Length 80;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6
|:|:|
```

```

DB 13 LLIWVP 18

RESULT 6
US-09-894-018-101
; Sequence 101, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV.PC3
US-09-894-018-101

Query Match          91.2%; Score 31; DB 10; Length 98;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6
|:|:|
DB 13 LLIWVP 18

RESULT 7
US-09-894-018-95
; Sequence 95, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 95
; LENGTH: 106
; TYPE: PRT
```

ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HCV.3s2  
US-09-894-018-95

Query Match 91.2%; Score 31; DB 10; Length 106;  
Best Local Similarity 66.7%; Pred. No. 58;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11:1:1  
DB 13 LLLWVP 18

RESULT 8  
US-09-894-018-93  
Sequence 93, Application US/09894018  
Patent No. US20020119127A1  
GENERAL INFORMATION:  
APPLICANT: EPIMUNE, Inc.  
APPLICANT: Sette, Alessandro  
APPLICANT: Chestnut, Robert  
APPLICANT: Livingston, Brian  
APPLICANT: Baker, Denisw  
APPLICANT: Newman, Mark  
APPLICANT: Brown, David

TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING  
TITLE OF INVENTION: MITIGENES AND PEPTIDES THEREBY  
FILE REFERENCE: 39963-20033.00  
CURRENT APPLICATION NUMBER: US/09/894, 018  
CURRENT FILING DATE: 2001-06-27  
PRIOR APPLICATION NUMBER: PCT/US00/35568  
PRIOR FILING DATE: 2000-12-28  
PRIOR FILING DATE: 1999-12-28  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: US 60/173,390  
PRIOR FILING DATE: 2001-04-16  
NUMBER OF SEQ ID NOS: 368  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 93  
LENGTH: 107  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HCV.3s1  
US-09-894-018-93

Query Match 91.2%; Score 31; DB 10; Length 107;  
Best Local Similarity 66.7%; Pred. No. 59;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11:1:1  
DB 13 LLLWVP 18

RESULT 9  
US-09-894-018-103  
Sequence 103, Application US/09894018  
Patent No. US20020119127A1  
GENERAL INFORMATION:  
APPLICANT: EPIMUNE, Inc.  
APPLICANT: Sette, Alessandro  
APPLICANT: Chestnut, Robert  
APPLICANT: Livingston, Brian  
APPLICANT: Baker, Denisw  
APPLICANT: Newman, Mark  
APPLICANT: Brown, David  
TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING  
TITLE OF INVENTION: MITIGENES AND PEPTIDES THEREBY  
FILE REFERENCE: 39963-20033.00  
CURRENT APPLICATION NUMBER: US/09/894, 018  
CURRENT FILING DATE: 2001-06-27

PRIOR APPLICATION NUMBER: PCT/US00/35568  
PRIOR FILING DATE: 2000-12-28  
PRIOR APPLICATION NUMBER: US 60/173,390  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: US 60/284,221  
PRIOR FILING DATE: 2001-04-16  
NUMBER OF SEQ ID NOS: 368  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 103  
LENGTH: 107  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HCV.PC4  
US-09-894-018-103

Query Match 91.2%; Score 31; DB 10; Length 107;  
Best Local Similarity 66.7%; Pred. No. 59;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11:1:1  
DB 13 LLLWVP 18

RESULT 10  
US-09-894-018-109  
Sequence 109, Application US/09894018  
Patent No. US20020119127A1  
GENERAL INFORMATION:  
APPLICANT: EPIMUNE, Inc.  
APPLICANT: Sette, Alessandro  
APPLICANT: Chestnut, Robert  
APPLICANT: Livingston, Brian  
APPLICANT: Baker, Denisw  
APPLICANT: Newman, Mark  
APPLICANT: Brown, David  
TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING  
TITLE OF INVENTION: MITIGENES AND PEPTIDES THEREBY  
FILE REFERENCE: 39963-20033.00  
CURRENT APPLICATION NUMBER: US/09/894, 018  
CURRENT FILING DATE: 2001-06-27  
PRIOR APPLICATION NUMBER: PCT/US00/35568  
PRIOR FILING DATE: 2000-12-28  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: US 60/173,390  
PRIOR FILING DATE: 2001-04-16  
NUMBER OF SEQ ID NOS: 368  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 109  
LENGTH: 123  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: AOSI.K  
US-09-894-018-109

Query Match 91.2%; Score 31; DB 10; Length 123;  
Best Local Similarity 66.7%; Pred. No. 67;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLIWIP 6  
11:1:1  
DB 13 LLLWVP 18

RESULT 11  
US-09-881-823-14  
Sequence 14, Application US/09881823  
Patent No. US20020068066A1  
GENERAL INFORMATION:  
APPLICANT: SHI, WENYUAN

APPLICANT: ANDERSON, MAXWELL  
APPLICANT: MORRISON, SHERIE  
APPLICANT: TRINH, RYAN  
APPLICANT: WIMS, LETITIA  
APPLICANT: CHEN, LI  
TITLE OF INVENTION: Method for the Treatment and Prevention of Dental Caries  
FILE REFERENCE: 22851-032  
CURRENT APPLICATION NUMBER: US/09/881,823  
CURRENT FILING DATE: 2001-06-15  
PRIOR APPLICATION NUMBER: US 07/378,577  
PRIOR FILING DATE: 1999-08-20  
NUMBER OF SEQ ID NOS: 32  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 14  
LENGTH: 128  
TYPE: PRT  
ORGANISM: Murine  
US-09-881-823-14

Query Match 91.2%; Score 31; DB 10; Length 128;  
Best Local Similarity 66.7%; Pred. No. 69;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
||:|:|  
Db 11 LLLWVP 16

RESULT 12  
US-09-992-524-6  
Sequence 6, Application US/09992524  
Patent No. US20020091240A1  
GENERAL INFORMATION:

APPLICANT: Vasquez, Maximiliano  
APPLICANT: Landolfi, Nicholas F.  
APPLICANT: Tsurushita, Naoya  
APPLICANT: Queen, Cary L.  
APPLICANT: Protein Design Labs, Inc.  
TITLE OF INVENTION: Humanized Antibodies To Gamma-Interferon  
FILE REFERENCE: 011823-008110US  
CURRENT APPLICATION NUMBER: US/09/992,524  
CURRENT FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: 09/450,520  
PRIOR FILING DATE: 1999-11-29  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 128  
TYPE: PRT  
ORGANISM: Artificial Sequence

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: human-mouse  
OTHER INFORMATION: transgenic construct HuZAF VL  
US-09-992-524-6

Query Match 91.2%; Score 31; DB 10; Length 128;  
Best Local Similarity 66.7%; Pred. No. 69;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
||:|:|  
Db 11 LLLWVP 16

RESULT 13  
US-09-894-018-99  
Sequence 99, Application US/09894018  
Patent No. US2002019127A1  
GENERAL INFORMATION:  
APPLICANT: EPIMUNE, Inc.  
APPLICANT: Sette, Alessandro  
APPLICANT: Chestnut, Robert  
APPLICANT: Livingston, Brian

APPLICANT: Baker, Dennis  
APPLICANT: Newman, Mark  
APPLICANT: Brown, David  
TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING  
MINISENS AND PEPTIDES THEREBY  
FILE REFERENCE: 39963-20033.00  
CURRENT APPLICATION NUMBER: US/09/894,018  
CURRENT FILING DATE: 2001-06-27  
PRIOR APPLICATION NUMBER: PCT/US00/35568  
PRIOR FILING DATE: 2000-12-28  
PRIOR APPLICATION NUMBER: US 60/173,390  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: US 60/284,221  
PRIOR FILING DATE: 2001-04-16  
NUMBER OF SEQ ID NOS: 368  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 99  
LENGTH: 130  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HCV.353  
US-09-894-018-99

Query Match 91.2%; Score 31; DB 10; Length 130;  
Best Local Similarity 66.7%; Pred. No. 70;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
||:|:|  
Db 13 LLLWVP 18

RESULT 14  
US-09-286-240-2  
Sequence 2, Application US/09286240  
Patent No. US20020010320A1  
GENERAL INFORMATION:  
APPLICANT: Felt, James W  
TITLE OF INVENTION: Chimeric and Humanized Antibodies to Angiogenin  
FILE REFERENCE: 10498/74073  
CURRENT APPLICATION NUMBER: US/09/286,240  
CURRENT FILING DATE: 1999-04-05  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 131  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-286-240-2

Query Match 91.2%; Score 31; DB 10; Length 131;  
Best Local Similarity 66.7%; Pred. No. 70;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
||:|:|  
Db 11 LLLWVP 16

RESULT 15  
US-09-879-461-2  
Sequence 2, Application US/09879461  
Publication No. US20020193575A1  
GENERAL INFORMATION:  
APPLICANT: Holmes, Stephen D.  
Gross, Mitchell S.  
Sylvester, Daniel R.  
TITLE OF INVENTION: Recombinant IL4 Antibodies Useful in  
Treatment of IL4 Mediated Disorders  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESS: SmithKline Beecham Corporation

STREET: Corporate Intellectual Property, UW2220 - 709  
Swedeland Rd.  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-2799  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/879,461  
FILING DATE: 12-Jun-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/612,929  
FILING DATE: <unknown>  
APPLICATION NUMBER: US 08/136,783  
FILING DATE: 14-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Sutton, Jeffrey A.  
REGISTRATION NUMBER: 34,028  
REFERENCE/DOCKET NUMBER: P50186-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 270-5024  
TELEFAX: (215) 270-5090  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 132 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-879-461-2

Query Match 91.2%; Score 31; DB 9; Length 132;  
Best Local Similarity 66.7%; Pred. NO. 71;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLIWIP 6  
||:|:|  
DB 11 LLLWVP 16

Search completed: January 3, 2003, 15:52:04  
Job time : 4.77273 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:21:45 ; Search time 26.5909 seconds  
(without alignments)  
30.067 Million cell updates/sec

Title: US-09-543-188a-4  
Perfect score: 44  
Sequence: 1 WLWMP 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.GeneSeq\_101002.\*

1: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:\*  
2: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:\*  
3: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:\*  
4: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:\*  
5: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:\*  
6: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:\*  
7: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:\*  
8: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:\*  
9: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:\*  
10: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:\*  
11: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:\*  
12: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:\*  
13: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:\*  
14: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:\*  
15: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:\*  
16: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:\*  
17: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:\*  
18: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:\*  
19: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:\*  
20: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:\*  
21: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:\*  
22: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:\*  
23: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	100.0	6	AAU11827	Peptide ligand for
2	42	99.5	173	ABB89811	Human polypeptide
3	40	90.9	137	AAW96472	Human reproductive
4	40	90.9	235	AAW41991	Human polypeptide
5	40	90.9	533	AAW40205	Human polypeptide
6	40	90.9	534	ABG30232	Novel human diapo
7	40	90.9	534	AAU32016	Novel human diapo
8	38	86.4	142	ABB67564	Drosophila melanog
9	38	86.4	402	AAU00427	Caenorhabditis ele
10	38	86.4	402	AAE22532	Caenorhabditis ele

11	37	84.1	27	AAU01575
12	37	84.1	384	ABP28377
13	37	84.1	385	ABB60764
14	37	84.1	393	AAU03712
15	37	84.1	415	AAU14114
16	37	84.1	415	AAU12993
17	37	84.1	458	ABP30101
18	37	84.1	461	ABB45843
19	37	84.1	561	AAW14440
20	37	84.1	836	AAW93820
21	37	84.1	2701	AAW99850
22	36	81.8	331	AAU46431
23	36	81.8	719	ABB47832
24	35	79.5	54	ABP31650
25	35	79.5	83	ABG37801
26	35	79.5	109	AAW87271
27	35	79.5	117	AAW89871
28	35	79.5	195	ABG10973
29	35	79.5	360	AAU35593
30	35	79.5	382	ABB76184
31	35	79.5	475	AAU34459
32	35	79.5	524	AAW74306
33	35	79.5	533	AAU33181
34	34	77.3	10	AAW32770
35	34	77.3	17	AAW16194
36	34	77.3	17	AAW68833
37	34	77.3	17	AAW58320
38	34	77.3	17	AAW09966
39	34	77.3	17	AAW17743
40	34	77.3	17	ABB72639
41	34	77.3	62	ABG01631
42	34	77.3	76	AAU50181
43	34	77.3	103	AAW24385
44	34	77.3	130	ABG01550
45	34	77.3	160	AAU43406

## ALIGNMENTS

RESULT 1  
AAU11827  
AAU11827 standard; peptide: 6 AA.  
XX  
AC AAU11827;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
Peptide ligand for Prion protein, PrP, #2.  
KW Prion protein; PrP; ligand; octapeptide motif; scrapie;  
KW Prion-associated disease; Creutzfeldt-Jakob disease;  
KW Gerstmann-Strausler-Scheinker disease; fatal familial insomnia;  
KW feline spongiform encephalopathy; bovine spongiform encephalopathy;  
KW transmissible mink encephalopathy; exotic ungulate encephalopathy;  
KW chronic wasting disease.  
XX  
OS Synthetic.  
XX  
PN WO200177687-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 05-APR-2001; 2001WO-US11150.  
XX  
PR 05-APR-2000; 2000US-0543188.  
XX  
PA (VITE-) VI TECHNOLOGIES INC.  
XX  
PI Hammond DJ, Wiltshire VR, Carbonei R, Shen H;  
XX WPI: 2002-061944/08.  
XX

Human secreted pro  
Streptococcus poly  
Drosophila melanog  
Group B Streptococ  
Non-receptor like  
Human type 3 Inost  
Propionibacterium  
Listeria monocytog  
Human ORF623 prote  
Human peptide enco  
Human signal pepti  
Human Immune/haema  
Novel human diapo  
Haemophilus influe  
Human G-protein co  
E. coli cellular p  
Neisseria meningit  
Novel human secret  
Human platelet gly  
Peptide containing  
Peptide binding in  
IL-1R1 binding pe  
Interleukin-1 type  
IL-1 antagonist pe  
Interleukin 1 anta  
Novel human diapo  
Propionibacterium  
Human EST encoded  
Novel human diapo  
Propionibacterium

PT New ligands for prion proteins, useful for detection or removal or  
PT prions and for treating prion-associated diseases, recognize a specific  
PT octapeptide motif -  
XX

PS Claim 16; Page 34; 47pp; English.

XX The invention relates to a ligand of less than 6 kD that binds to a  
CC polypeptide containing the sequence glyTyrGlylnProHisGlyGly (A) or an  
CC analogue that is the retro-inverso isomer of (A). The sequence A is  
CC an octapeptide motif from the prion protein (PrP). The ligands are  
CC identified by binding assays with the peptide (A) or peptides containing  
CC (A). The ligands are used for detecting prion proteins (or prions) in  
CC biological or environmental samples, e.g. for diagnosis, also for  
CC removing them from samples, for treating or retarding development of  
CC prion-associated diseases, especially Creutzfeldt-Jakob diseases (in  
CC iatrogenic, new variant, familial or sporadic forms), but also  
CC Gerstmann-Strussler-Scheinker disease, fatal familial insomnia,  
CC scrapie, bovine or feline spongiform encephalopathy, transmissible mink  
CC or exotic ungulate encephalopathy, or chronic wasting disease. The  
CC present sequence is a ligand of the invention.

Sequence 6 AA:

Query Match 100.0%; Score 44; DB 23; Length 6;  
Best Local Similarity 100.0%; Pred. NO. 7.8e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLXWIP 6  
DB 1 WLXWIP 6

RESULT 2  
ABB89811

ID ABB89811 standard; Protein; 173 AA.

AC ABB89811;

DT 24-MAY-2002 (first entry)

DE Human polypeptide SEQ ID NO 2187.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein.

XX Homo sapiens.

PN WO200190304-A2.

PD 29-NOV-2001.

PF 18-MAY-2001; 2001WO-US16450.

PR 19-MAY-2000; 2000US-205515P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Blirze CE, Rosen CA;

DR WPI; 2002-122018/16.

DR N-PSDB; ABL90220.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
PT prevention of neural, immune system, muscular, reproductive,  
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
PT disorders -

PS Claim 11; SEQ ID NO 2187; 2081pp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABL90853) and proteins

CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 173 AA:

Query Match 95.5%; Score 42; DB 23; Length 173;  
Best Local Similarity 83.3%; Pred. NO. 13;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLXWIP 6  
DB 15 WLXWIP 20

RESULT 3  
AAM96472

ID AAM96472 standard; Protein; 137 AA.

AC AAM96472;

DT 21-NOV-2001 (first entry)

DE Human reproductive system related antigen SEQ ID NO: 5130.

XX Human; reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy.

XX Homo sapiens.

PN WO200155320-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01339.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224516.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.





AC AAM41991;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 6922.  
XX  
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
XX Homo sapiens.  
OS  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
XX 26-DEC-2000; 2000WO-US34263.  
XX  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Dmanac RT;  
XX  
DR WPI: 2001-442253/47.  
DR N-PSDB: AAI61147.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
XX Example 2; SEQ ID NO 6922; 10078pp; English.  
PS  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
XX Sequence 235 AA:  
SQ  
Query Match 90.9%; Score 40; DB 22; Length 235;  
Best Local Similarity 66.7%; Pred. No. 37;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

AAM40205  
ID AAM40205 standard; Protein: 533 AA.  
XX  
AC AAM40205;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 3350.  
XX  
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX  
XX Homo sapiens.  
OS  
PN WO200153312-A1.  
XX  
PD 26-JUL-2001.  
XX  
XX 26-DEC-2000; 2000WO-US34263.  
XX  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Dmanac RT;  
XX  
DR WPI: 2001-442253/47.  
DR N-PSDB: AAI59361.  
XX  
PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries -  
XX  
XX Example 5; SEQ ID NO 3350; 10078pp; English.  
PS  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
XX Sequence 533 AA:  
SQ  
Query Match 90.9%; Score 40; DB 22; Length 533;  
Best Local Similarity 66.7%; Pred. No. 84;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

## RESULT 6

ABG30232 standard; Protein: 534 AA.

ABG30232;

18-FEB-2002 (first entry)

Novel human diagnostic protein #30223.

Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.

Homo sapiens.

MO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

N-PSDB; AAS94419.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity

Claim 20; SEQ ID No 60591; 103bp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 534 AA;

Query Match 90.9%; Score 40; DB 22; Length 534;

Best Local Similarity 66.7%; Pred. No. 84;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWYIP 6

Db 511 WYWYWP 516

## RESULT 7

AAU32016 standard; Protein: 534 AA.

AAU32016;

18-DEC-2001 (first entry)

Novel human secreted protein #2507.

Human; vaccination; gene therapy; nutritional supplement; stem cell proliferation; haematopoiesis; nerve tissue regeneration; immune suppression; immune stimulation; anti-inflammatory; leukaemia.

Homo sapiens.

WO200179449-A2.

25-OCT-2001.

16-APR-2001; 2001WO-US08656.

18-APR-2000; 2000US-0552929.

26-JAN-2001; 2001US-0770160.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-611725/70.

Nucleic acids encoding a range of human polypeptides, useful in genetic vaccination, testing and therapy -

Claim 20; Page 541; 765bp; English.

The invention relates to novel human secreted polypeptides. The polypeptides and antibodies to the polypeptides are useful for determining the presence of or predisposition to a disease associated with altered levels of polypeptide. The polypeptides are also useful for identifying agents (agonists and antagonists) that bind to them. Cells expressing the proteins are useful for identifying a therapeutic agent for use in treatment of a pathology related to aberrant expression or physiological interactions of the polypeptide. Vectors comprising the nucleic acids encoding the polypeptides and cells genetically engineered to express them are also useful for producing the proteins. The proteins are useful in genetic vaccination, testing and therapy, and can be used as nutritional supplements. They may be used to increase stem cell proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon and/or nerve tissue growth or regeneration; immune suppression and/or stimulation; as anti-inflammatory agents; and in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid sequences of novel human secreted proteins of the invention.

Sequence 534 AA;

Query Match 90.9%; Score 40; DB 22; Length 534;

Best Local Similarity 66.7%; Pred. No. 84;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWYIP 6

Db 511 WYWYWP 516

## RESULT 8

ABB67564 standard; Protein: 142 AA.

ABB67564;

26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 29484.  
 XX  
 XX Drosophila: developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PP 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 DK WPI: 2001-655660/75.  
 XX N-PSDB: ABL11667.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure: SEQ ID NO 29484; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB13511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins  
 CC (AAB57737-AB872072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 142 AA:  
 Query Match 86.4%; Score 38; DB 22; Length 142;  
 Best Local Similarity 83.3%; Pred. NO. 48;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 1 WLYWIP 6  
 1 1111  
 53 WRYWIP 58

RESULT 9  
 AAU00427  
 ID AAU00427 standard; Protein; 402 AA.  
 XX  
 AC AAU00427;  
 XX  
 DT 11-MAY-2001 (first entry)  
 XX  
 DE Caenorhabditis elegans omega-3 fatty acyl desaturase fat-1 polypeptide.  
 XX  
 KW Omega-3 fatty acyl desaturase; fat-1; omega-6 fatty acid; lipid;  
 KW omega-3 fatty acid; omega-3 desaturase activity; food; oil;  
 KW nutritional supplement; chemical feedstock.  
 XX  
 OS Caenorhabditis elegans.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 80..124  
 FT /note- "Region containing stretches of hydrophobic  
 FT residues similar to those found in other  
 FT desaturases"

FT Misc-feature 123  
 FT /note- "Highly conserved in membrane desaturases"  
 FT Misc-feature 127  
 FT /note- "Highly conserved in membrane desaturases"  
 FT Misc-feature 159  
 FT /note- "Highly conserved in membrane desaturases"  
 FT Misc-feature 162..163  
 FT /note- "Highly conserved in membrane desaturases"  
 FT Region 229..284  
 FT /note- "Region containing stretches of hydrophobic  
 FT residues similar to those found in other  
 FT desaturases"  
 FT Misc-feature 324  
 FT /note- "Highly conserved in membrane desaturases"  
 FT Misc-feature 327..328  
 FT /note- "Highly conserved in membrane desaturases"  
 PN US6194167-B1.  
 PD 27-FEB-2001.  
 PP 18-FEB-1998; 98US-0025578.  
 PR 18-FEB-1997; 97US-0038409.  
 PA (UNIV ) UNIV WASHINGTON STATE RES FOUND.  
 PI Browse JA, Szychalla JP;  
 DK WPI: 2001-217927/22.  
 XX N-PSDB: AAS00431.  
 PT Novel host cell transformed with recombinant fat-1 polynucleotide  
 PT encoding polypeptide which desaturates omega-6 fatty acid to  
 PT corresponding omega-3 fatty acid, useful for producing lipids -  
 XX  
 PS Claim 1; Fig 1; 17pp; English.  
 XX  
 CC The present sequence represents Caenorhabditis elegans omega-3 fatty  
 CC acyl desaturase fat-1 polypeptide. The polypeptide desaturates an  
 CC omega-6 fatty acid to a corresponding omega-3 fatty acid by  
 CC catalysing the introduction of an omega-3 double bond into 18-, 20-, and  
 CC 22- carbon fatty acids. Recombinant expression of C. elegans fat-1  
 CC polynucleotide in a wide variety of host cells, including Arabidopsis  
 CC thaliana and Saccharomyces cerevisiae, produces a polypeptide with  
 CC omega-3 desaturase activity. A host cell transformed with C. elegans  
 CC fat-1 polynucleotide is useful for desaturating an omega-6 fatty acid  
 CC to an omega-3 fatty acid by growing the host cell under conditions under  
 CC which the polypeptide is expressed. The fat-1 polypeptide is useful for  
 CC producing lipids having a higher proportion of omega-3 fatty acid and  
 CC the lipids are useful as food, oils, as nutritional supplements, and as  
 CC chemical feedstocks.  
 XX  
 SQ Sequence 402 AA:  
 Query Match 86.4%; Score 38; DB 22; Length 402;  
 Best Local Similarity 66.7%; Pred. NO. 1.3e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 WLYWIP 6  
 1 11:1  
 DB 261 WRYWIP 266

RESULT 10  
 AAE22532  
 ID AAE22532 standard; Protein; 402 AA.  
 XX  
 AC AAE22532;  
 XX  
 DT 26-JUL-2002 (first entry)  
 XX  
 DE Caenorhabditis elegans fat-1 (omega-3 fatty acyl desaturase) protein.

```
XX FAT-1 protein; omega-3 fatty acyl desaturase; chemical feedstock;
KM food oil; nutritional supplement; enzyme.
XX
XX Caenorhabditis elegans.
XX
XX US2002042933-A1.
XX
XX 11-APR-2002.
XX
XX 20-DEC-2000; 2000US-0747755.
XX
XX 18-FEB-1997; 97US-038409P.
XX
XX 18-FEB-1998; 98US-0025578.
XX
XX (BROWSE J A.
XX (SPYC/) SPYCHALLA J P.
XX
XX Browse JA, Spychalla JP;
XX
XX MPI: 2002-361389/39.
XX N-PSDB; AAD35559.
XX
XX Recombinant expression of fat-1 gene of Caenorhabditis elegans in a
XX wide variety of cells, including cells of Arabidopsis thaliana and
XX Saccharomyces cerevisiae, produces a polypeptide having omega-3
XX desaturase activity -
XX
XX Example 1; Fig 1; 20pp; English.
XX
XX The present sequence is Caenorhabditis elegans FAT-1, an omega-3 fatty
XX acyl desaturase protein. FAT-1 polypeptide is useful for producing lipids
XX having a higher proportion of omega-3 fatty acid. Such lipids are useful
XX as food oils nutritional supplements and as chemical feedstocks.
XX
XX SQ Sequence 402 AA;
XX
XX Query Match 86.4%; Score 38; DB 23; Length 402;
XX Best Local Similarity 66.7%; Pred. No. 1.3e+02;
XX Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 WLXWIP 6
XX 11111
XX Db 261 WXYWVP 266
XX
XX RESULT 11
XX A01575
XX AAU01575 standard; Peptide; 27 AA.
XX
XX AC AAU01575;
XX
XX 18-JUL-2001 (first entry)
XX
XX Human secreted protein immunogenic epitope encoded by gene #15.
XX
XX Human secreted protein; autoimmune disorder; hyperproliferative disorder;
XX cardiovascular disorder; cerebrovascular disorder; angiogenesis;
XX nervous system disorder; bacterial infection; viral infection;
XX fungal infection; ocular disorder; wound healing; tissue regeneration;
XX epithelial cell proliferation; skin ageing; chemotaxis; IgG Fc region.
XX
XX Homo sapiens.
XX
XX WO200123547-A1.
XX
XX 05-APR-2001.
XX
XX 26-SEP-2000; 2000WO-US26337.
XX
XX 27-SEP-1999; 99US-0155806.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
```

```
XX Komatsoulis GA, Ruben SM, Rosen CA;
XX
XX MPI: 2001-266151/27.
XX
XX N-PSDB; AAS02654.
XX
XX Nucleic acids encoding 26 human secreted polypeptides, useful for
XX preventing, diagnosing and/or treating e.g. Gaucher's disease,
XX Alzheimer's disease, Schmitzer syndrome, Creutzfeldt-Jacob disease,
XX diabetes mellitus and multiple sclerosis -
XX
XX Disclosure; Page 387; 412pp; English.
XX
XX The sequence represents a human secreted protein encoded by a nucleic
XX acid of the invention. Secreted proteins and their related nucleic acids
XX can be used in the diagnosis of or susceptibility to a pathological
XX condition by determining the presence or absence of a mutation in a
XX nucleic acid or the presence or amount of expression of a secreted
XX protein. The sequences are used to prevent, treat or ameliorate a medical
XX condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
XX chickens or sheep. The antibodies to the polypeptides can also be used in
XX alleviating symptoms associated with disorders and in diagnostic
XX immunoassays e.g. radioimmunoassays or enzyme linkedCC immunosorbent
XX assays (ELISA). The disorders include autoimmune diseases e.g. rheumatoid
XX arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
XX liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
XX disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders
XX e.g. Alzheimer's disease, infections caused by bacteria, viruses and
XX fungi and ocular disorders e.g. corneal infection. The peptides can also
XX be used to aid wound healing and epithelial cell proliferation, to help
XX prevent skin ageing due to sunburn, to maintain organs before
XX transplantation, to regenerate tissues, in chemotaxis and as a food
XX additive or preservative to alter storage capabilities.
XX
XX SQ Sequence 27 AA;
XX
XX Query Match 84.1%; Score 37; DB 22; Length 27;
XX Best Local Similarity 66.7%; Pred. No. 13;
XX Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 WLXWIP 6
XX 11111
XX Db 8 WLWMLP 13
XX
XX RESULT 12
XX ABP28377
XX ID ABP28377 standard; Protein; 384 AA.
XX
XX AC ABP28377;
XX
XX 02-JUL-2002 (first entry)
XX
XX Streptococcus polypeptide SEQ ID NO 5930.
XX
XX Streptococcus polypeptide
XX
XX Streptococcus: GAS; GBS: group B streptococcus; Streptococcus agalactiae;
XX group A streptococcus; Streptococcus pyogenes; antibacterial;
XX antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
XX Streptococcus agalactiae.
XX
XX MO200234771-A2.
XX
XX 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-GB04789.
XX
XX 27-OCT-2000; 2000GB-0026333.
XX
XX 24-NOV-2000; 2000GB-0028727.
XX
XX 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.
```



CC useful in the preparation of a medicament for the treatment or  
 CC prophylaxis of Group B Streptococcus infection. The invention does not  
 CC have the disadvantages of varied response rate associated with prior art  
 CC capsid polysaccharide vaccination against Group B Streptococcus.

XX Sequence 393 AA;

Query Match 84.1%; Score 37; DB 22; Length 393;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLYWI 5  
 |||||  
 Db 99 WLYWI 103

# RESULT 15

AAR14114  
 ID AAR14114 standard; Protein; 415 AA.

AAR14114;

DT 09-DEC-1991 (first entry)

DE Non-receptor linked protein tyrosine phosphatase.

KW PTase; malignancy; cancer.

OS Homo sapiens.

XX Key Location/Qualifiers

FH Region 42..274

FT Active-site 377..381

FT Domain /label= nuclear recognition signal

FT 396..415

FT /label= C-terminal

FT /note= "hydrophobic"

PN WO9113989-A.

PD 19-SEP-1991.

PE 14-MAR-1991; 91WO-US01748.

PR 14-MAR-1990; 90US-0494036.

XX (WASH-) WASHINGTON RES FOUN.

XX Fischer EH, Krebs EG, Tonks NK, Cool DE;

DR WPI: 1991-295643/40.

DR N-PSDB: AAQ13802.

XX New DNA encoding non-receptor-linked protein tyrosine phosphatase

PT - used in treating malignancies associated with protein-linked

PT tyrosine phosphatase.

XX Disclosure: Fig 1; 38pp; English.

XX The sequence was deduced from a cDNA isolated from a cDNA library

CC prepd. from peripheral T cell poly(A)+ mRNA. The gene can be

CC used to express recombinant PTase or derivs. useful in cancer

CC therapy where a protein kinase is involved.

XX Sequence 415 AA;

Query Match 84.1%; Score 37; DB 12; Length 415;

Best Local Similarity 83.3%; Pred. No. 2e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 382 WLYWQP 387

Search completed: January 3, 2003, 15:28:37  
 JOD time : 28.5909 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:00 : Search time 9.95455 Seconds  
(without alignments)  
57.944 Million cell updates/sec

Title: US-09-543-188a-4

Perfect score: 44

Sequence: 1 WLWYIP 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	95.5	3944	2 T19997	hypothetical prote
2	40	90.9	533	2 T00742	ubiquitin-binding
3	39	88.6	663	1 B54075	arachidonate 12-11
4	39	88.6	663	1 S30051	arachidonate 12-11
5	39	88.6	663	2 I52462	arachidonate 12-11
6	39	88.6	1564	2 T27121	hypothetical prote
7	38	86.4	211	2 C83360	hypothetical prote
8	38	86.4	407	2 T48817	hypothetical prote
9	38	86.4	739	2 T25030	hypothetical prote
10	37	84.1	415	1 A33899	protein-tyrosine-P
11	37	84.1	461	2 AF1079	amino acid transpo
12	37	84.1	525	2 T41427	membrane transport
13	37	84.1	938	2 S55051	Bicaudal-C - fruit
14	37	84.1	1613	2 JE0272	low density lipopr
15	37	84.1	1613	2 JE0273	low density lipopr
16	37	84.1	2701	2 S17796	inositol-trisphosp
17	36	81.8	298	2 C90880	probable phosphat
18	36	81.8	298	2 F85738	probable phosphat
19	36	81.8	340	2 A75575	probable homoprot
20	36	81.8	456	2 S47924	Mgal protein - yea
21	36	81.8	461	1 G64537	2-oxoglutarate/mal
22	36	81.8	564	2 T15477	hypothetical prote
23	36	81.8	719	2 AE1131	hypothetical prote
24	35	79.5	55	2 E82259	hypothetical prote
25	35	79.5	126	2 AG1932	hypothetical prote
26	35	79.5	136	2 AG1517	hypothetical prote
27	35	79.5	152	2 S50967	probable membrane
28	35	79.5	255	2 AE0537	hypothetical prote
29	35	79.5	264	2 S78700	probable export pr

30	35	79.5	264	2 AF0753	flagellar biosynth
31	35	79.5	269	2 T10253	membrane protein M
32	35	79.5	273	2 B83551	hypothetical prote
33	35	79.5	279	2 T10251	membrane protein M
34	35	79.5	300	2 F36790	hypothetical prote
35	35	79.5	305	2 S58822	probable membrane
36	35	79.5	360	1 A64185	phospho-N-acetylmu
37	35	79.5	371	2 H97452	cytochrome-c oxida
38	35	79.5	366	2 T25699	hypothetical prote
39	35	79.5	467	2 S49268	probable arginine
40	35	79.5	470	2 A69751	histidine permease
41	35	79.5	475	2 D64751	amino acid permeas
42	35	79.5	528	2 S26948	cytochrome-c oxida
43	35	79.5	538	2 S36424	cytochrome-c oxida
44	35	79.5	541	1 OD231	cytochrome-c oxida
45	35	79.5	541	2 A48327	cytochrome-c oxida

#### ALIGNMENTS

RESULT 1  
T19997  
hypothetical protein C47D12.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T19997  
R:Gajadsky, S.  
Submitted to the EMBL Data Library, March 1996  
A:Reference number: Z19209  
A:Accession: T19997  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-3944 <WIL>  
A:Cross-references: EMBL:Z69902; PIDN:CAA93765.1; GSPDB:GN00020; CESP:C47D12.1  
A:Experimental source: clone C47D12  
C:Genetics:  
A:Gene: CESP:C47D12.1  
A:Map position: 2  
A:Introns: 46/3; 308/3; 408/2; 449/3; 820/3; 946/3; 1069/1; 1634/3; 1737/2; 2009/2; 2

Query Match  
Best Local Similarity 95.5%; Score 42; DB 2; Length 3944;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWYIP 6  
|||||  
Db 3252 WLWYIP 3257

RESULT 2  
T00742  
ubiquitin-binding protein homolog A-735G6.2 - human  
C:Species: Homo sapiens (man)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001  
C:Accession: T00742  
R:Adams, M.D.; Loftus, B.J.; Zhou, L.; Crosby, M.; Fuhrmann, J.; Brandon, R.; Kim, U.  
Submitted to the EMBL Data Library, March 1998  
A:Description: Human Chromosome 16 BAC clone C1987SK-A-735G6.  
A:Reference number: Z14181  
A:Accession: T00742  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-533 <ADA>  
A:Cross-references: EMBL:AC002400; NID:92576344; PIDN:AAC05812.1; PID:92576346  
C:Genetics:  
A:Gene: A-735G6.2  
A:Map position: 16p12  
A:Introns: 233/1; 343/1; 412/3; 434/3; 470/1; 497/3  
A:Note: A-735G6.2

Query Match  
Best Local Similarity 90.9%; Score 40; DB 2; Length 533;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLYWIP 6  
11111  
Db 510 WLYWIP 515

## RESULT 3

B54075  
archidonate 12-11poxigenase (EC 1.13.11.31), leukocyte [validated] - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: B54075; 149439  
R:Chen, X.S.; Kure, U.; Jenkins, N.A.; Copeland, N.G.; Funk, C.D.  
J. Biol. Chem. 269, 13979-13987, 1994  
A>Title: CDNA cloning, expression, mutagenesis of C-terminal isoleucine, genomic structure  
A:Reference number: A54075; MUID:94245713; PMID:8188678  
A:Accession: B54075  
A:Molecule type: DNA; mRNA  
A:Residues: 1-663 <RES>  
A:Cross-references: GB:U04331; NID:9467216; PIDN:AAA20658.1; PID:9467217  
A:Experimental source: strains C57BL/6 and ICR, spleen leukocytes  
A>Note: removal or substitution of ile-663 abolished enzyme activity  
R:Freire-Moar, J.; Alavi-Nassab, A.; Ng, M.; Mulkins, M.; Sigal, E.  
Biochim. Biophys. Acta 1254, 112-116, 1995  
A>Title: Cloning and characterization of a murine macrophage 11poxigenase.  
A:Reference number: 149439; MUID:95110857; PMID:76111740  
A:Accession: 149439  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-36; 'N', 38-118, 'O', 120-396, 'N', 398-663 <RES>  
A:Cross-references: GB:L34570; NID:9509607; PIDN:AAA64930.1; PID:9763530  
C:Comment: A second archidonate 12-11poxigenase from mouse platelets is shown in (PIR-A  
C:Genetics:  
A:Map position: 11  
C:Superfamily: archidonate 5-11poxigenase  
C:Keywords: oxidoreductase

Query Match 88.6%; Score 39; DB 1; Length 663;  
Best Local Similarity 66.7%; Pred. No. 68;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLYWIP 6  
11111  
Db 551 WLYWIP 556

## RESULT 4

B5051  
archidonate 12-11poxigenase (EC 1.13.11.31) - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S30051  
R:Matanabe, T.; Medina, J.F.; Haegstroom, J.Z.; Radmark, O.; Samuelsson, B.  
Eur. J. Biochem. 212, 605-612, 1993  
A>Title: Molecular cloning of a 12-11poxigenase cDNA from rat brain.  
A:Reference number: S30051; MUID:93185682; PMID:8444196  
A:Accession: S30051  
A:Molecule type: mRNA  
A:Residues: 1-663 <RES>  
A:Cross-references: EMBL:L06040; NID:9205212; PIDN:AAA41532.1; PID:9205213  
C:Superfamily: archidonate 5-11poxigenase  
C:Keywords: oxidoreductase

Query Match 88.6%; Score 39; DB 1; Length 663;  
Best Local Similarity 66.7%; Pred. No. 68;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLYWIP 6  
11111  
Db 551 WLYWIP 556

RESULT 5  
152462  
archidonate 12-11poxigenase (EC 1.13.11.31) - rat  
C:Species: Rattus sp. (rat)  
C>Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 11-Jun-1999  
C:Accession: 152462  
R:Hada, T.; Hagiya, H.; Suzuki, H.; Arakawa, T.; Nakamura, M.; Matsuda, S.; Yoshimoto  
Biochim. Biophys. Acta 1211, 221-228, 1994  
A>Title: Archidonate 12-11poxigenase of rat pineal glands: catalytic properties and  
A:Reference number: 152462; MUID:94162305; PMID:8117750  
A:Accession: 152462  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-663 <RES>  
A:Cross-references: GB:S69383; NID:9545793; PIDN:AAB30132.1; PID:9545794  
C:Superfamily: archidonate 5-11poxigenase  
C:Keywords: oxidoreductase

Query Match 88.6%; Score 39; DB 2; Length 663;  
Best Local Similarity 66.7%; Pred. No. 68;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLYWIP 6  
11111  
Db 551 WLYWIP 556

## RESULT 6

T27121  
hypothetical protein Y53C10A.9 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
C:Accession: T27121  
R:White, S.  
Submitted to the EMBL Data Library, November 1998  
A:Reference number: Z20314  
A:Accession: T27121  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1564 <MUI>  
A:Cross-references: EMBL:AL033536; PIDN:CAA22142.1; CESP:Y53C10A.9  
A:Experimental source: clone Y53C10A  
C:Genetics:  
A:Gene: CESP:Y53C10A.9  
A:Introns: 43/3; 92/2; 148/2; 226/3; 354/1; 712/3; 817/1; 875/1; 916/3; 984/3; 1069/2  
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Query Match 88.6%; Score 39; DB 2; Length 1564;  
Best Local Similarity 66.7%; Pred. No. 1,564; 02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLYWIP 6  
11111  
Db 326 WLYWIP 331

## RESULT 7

C83360  
hypothetical protein PA2287 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: C83360  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; L.  
Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: C83360  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-211 <STO>  
A:Cross-references: GB:AE004654; GB:AE004091; NID:99948311; PIDN:AAG05675.1; GSPDB:GN

A:Experimental source: strain PA01  
C:Gene: PA2287

Query Match 86.4%; Score 38; DB 2; Length 211;  
Best Local Similarity 66.7%; Pred. No. 33;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLYWIP 6  
|||:|  
DB 59 WLFWIP 64

#### RESULT 8

T48817  
hypothetical protein 68B2.10 [imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 19-May-2000  
C:Accession: T48817  
Schulte, U.; Aign, V.; Hohnsels, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, R.  
Submitted to the Protein Sequence Database, April 2000  
Reference number: Z24541

A:Accession: T48817  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-407 <SCH>  
A:Cross-references: EMBL:AL353821; GSPDB:GN00112; NCSP:68B2.10  
A:Experimental source: cosmid contig 68B2; strain 74  
C:Genetics:

A:Gene: NCSP:68B2.10  
A:Map position: 2  
A:Introns: 56/2; 68/3  
C:Superfamily: Neurospora crassa hypothetical protein 68B2.10

Query Match 86.4%; Score 38; DB 2; Length 407;  
Best Local Similarity 66.7%; Pred. No. 62;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLYWIP 6  
|||:|  
DB 172 WLFWIP 177

#### RESULT 9

T25030  
hypothetical protein T20D3.11 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
C:Accession: T25030  
Alloyd, C.  
Submitted to the EMBL Data Library, December 1995  
Reference number: Z19971

A:Accession: T25030  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-739 <WIL>  
A:Cross-references: EMBL:Z68220; PIDN:CA92491.1; GSPDB:GN00022; CESP:T20D3.11  
A:Experimental source: clone T20D3  
C:Genetics:

A:Gene: CESP:T20D3.11  
A:Map position: 4  
A:Introns: 205/1; 246/3; 437/2; 472/1; 643/3; 721/2  
C:Superfamily: Caenorhabditis elegans hypothetical protein T20D3.11

Query Match 86.4%; Score 38; DB 2; Length 739;  
Best Local Similarity 83.3%; Pred. No. 11e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLYWIP 6  
|||:|  
DB 475 WSYWIP 480

#### RESULT 10

A33899  
protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 2 - human  
N:Alternate names: protein-tyrosine-phosphatase, T-cell specific  
C:Species: Homo sapiens (man)  
C>Date: 31-Mar-1990 #sequence\_revision 02-May-1994 #text\_change 07-May-1999  
C:Accession: A33899; C60345; A45742

R:COOL, D.E.; Tonks, N.K.; Charbonneau, H.; Walsh, K.A.; Fischer, E.H.; Krebs, E.G.  
Proc. Natl. Acad. Sci. U.S.A. 86: 5257-5261, 1989  
A:Title: cDNA isolated from a human T-cell library encodes a member of the protein-ty  
A:Reference number: A33899; MUID:89315776; PMID:2546150  
A:Accession: A33899  
A:Molecule type: mRNA  
A:Residues: 1-415 <COO>

A:Cross-references: GB:M25393  
R:Champion-Arnaud, P.; Gesnel, M.C.; Foulkes, N.; Ronsin, C.; Sassone-Corsi, P.; Brea  
Oncogene 6, 1203-1209, 1991  
A:Title: Activation of transcription via AP-1 or CREB regulatory sites is blocked by  
A:Reference number: A60345; MUID:91319401; PMID:1650442  
A:Accession: C60345

A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 348-415 <CHA>

R:Johnson, C.V.; COOL, D.E.; Glaccum, M.B.; Green, N.; Fischer, E.H.; Bruskun, A.; Hi  
Genomics 16, 619-629, 1993  
A:Title: Isolation and mapping of human T-cell protein tyrosine phosphatase sequences  
Obs.

A:Reference number: A45742; MUID:93315152; PMID:8325634  
A:Accession: A45742  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: DNA

A:Residues: 115-125; 160-170; 230-241; 413-415 <OH>  
C:Comment: This protein and protein-tyrosine-phosphatase 11A are produced from the sa  
C:Genetics:

A:Gene: GDB:PTPN2; PTP  
A:Cross-references: GDB:128098; OMIM:176887  
A:Map position: 18p11.22-18p11.21  
A:Introns: 120/3; 165/3; 235/3; 347/2

A>Note: list of introns may be incomplete  
C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 1; protein-tyrosine-ph  
C:Keywords: alternative splicing; phosphoprotein; phosphoric monoester hydrolase; tyr

F:42-264/Domain: protein-tyrosine-phosphatase homology <PTP>  
F:216/Active site: Cys (phosphocysteine intermediate) #status predicted  
F:222/Binding site: substrate phosphate (Arg) #status predicted

Query Match 84.1%; Score 37; DB 1; Length 415;  
Best Local Similarity 83.3%; Pred. No. 91;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLYWIP 6  
|||:|  
DB 382 WLYWIP 387

#### RESULT 11

AF1079  
amino acid transporter homolog lmo0037 [imported] - Listeria monocytogenes (strain EG  
C:Species: Listeria monocytogenes  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AF1079

R:Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maltournam, A.;  
Ok, C.; Schueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AF1079

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-461 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC98252.1; PID:G16409396; GSPDB:GN00177

A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo0037

Query Match 84.1%; Score 37; DB 2; Length 461;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLXWI 5  
|||||  
DB 86 WLXWI 90

RESULT 12  
T41427

membrane transporter - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C:Accession: T41427

Pod, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.  
Submitted to the EMBL Data Library, September 1998  
Reference number: 221954

A:Accession: T41427  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-525 <MOO>

A:Cross-references: EMBL:AL031798; PIDN:CAA21196.1; GSPDB:GN00068; SPDB:SPCC576.17c  
A:Experimental source: strain 972h; cosmid c576  
C:Genetics:

A:Gene: SPDB:SPCC576.17c  
A:Map position: 3  
A:Introns: 19/3

C:Superfamily: amiloride resistance protein carl

Query Match 84.1%; Score 37; DB 2; Length 525;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLXWI 5  
|||||  
DB 238 WLXWI 242

RESULT 13  
S55051

Bicaudal-C - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster  
C>Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 29-May-1998  
C:Accession: S55051

ahone, M.; Saffman, E.E.; Lasko, P.F.  
EMBO J. 14, 2043-2055, 1995

A:Title: Localized Bicaudal-C RNA encodes a protein containing a KH domain, the RNA bind

A:Reference number: S55051; MUID:95262644; PMID:7538070

A:Accession: S55051  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-938 <MAH>

A:Cross-references: EMBL:U15928  
C:Genetics:

A:Gene: FlyBase:Bicc  
A:Cross-references: FlyBase:FBgn0000182

A:Introns: 118/3; 142/1; 168/3; 277/3; 319/3; 447/1; 822/1

C:Superfamily: SAM homology

F:800-865/Domain: SAM homology <SAM>

Query Match 84.1%; Score 37; DB 2; Length 938;  
Best Local Similarity 66.7%; Pred. No. 2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLXWI 6  
|||||  
DB 30 WLXWI 35

RESULT 14  
JE0272

low density lipoprotein receptor-related protein 6 - human

C:Species: Homo sapiens (man)  
C>Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 02-Aug-2002  
C:Accession: JE0272

R:Brown, S.D.; Twells, R.C.J.; Hey, P.J.; Cox, R.D.; Levy, E.R.; Soderman, A.R.; Metz

Biochem. Biophys. Res. Commun. 248, 879-888, 1998

A:Title: Isolation and characterization of LRP6, a novel member of the low density li

A:Reference number: JE0272; MUID:98369644; PMID:9704021

A:Accession: JE0272  
A:Molecule type: mRNA

A:Residues: 1-1613 <BRO>

A:Cross-references: GB:AF074264; NID:93462526; PIDN:AC33006.1; PID:93462527

C:Genetics:  
A:Gene: LRP6

A:Map position: 12p11-12p13

C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL re

F:286-323/Domain: EGF homology <EGF1>

F:592-627/Domain: EGF homology <EGF>

F:1207-1243/Domain: EGF homology <EGF2>

F:1249-1285/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F:1288-1322/Domain: LDL receptor ligand-binding repeat homology <LDL3>

Query Match 84.1%; Score 37; DB 2; Length 1613;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLXWI 5  
|||||  
DB 1156 WLXWI 1160

RESULT 15  
JE0273

low density lipoprotein receptor-related protein 6 - mouse

C:Species: Mus musculus (house mouse)  
C>Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 02-Aug-2002  
C:Accession: JE0273

R:Brown, S.D.; Twells, R.C.J.; Hey, P.J.; Cox, R.D.; Levy, E.R.; Soderman, A.R.; Metz

Biochem. Biophys. Res. Commun. 248, 879-888, 1998

A:Title: Isolation and characterization of LRP6, a novel member of the low density li

A:Reference number: JE0272; MUID:98369644; PMID:9704021

A:Accession: JE0273  
A:Molecule type: mRNA

A:Residues: 1-1613 <BRO>

A:Cross-references: GB:AF074265; NID:93462528; PIDN:AC33007.1; PID:93462529

C:Genetics:  
A:Gene: Lrp6

A:Map position: 6

C:Superfamily: low density lipoprotein receptor-related protein; EGF homology; LDL re

F:286-323/Domain: EGF homology <EGF1>

F:592-627/Domain: EGF homology <EGF>

F:1207-1243/Domain: EGF homology <EGF2>

F:1249-1285/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F:1288-1322/Domain: LDL receptor ligand-binding repeat homology <LDL3>

F:1326-1360/Domain: LDL receptor ligand-binding repeat homology <LDL3>

Query Match 84.1%; Score 37; DB 2; Length 1613;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLXWI 5  
|||||  
DB 1156 WLXWI 1160

Search completed: January 3, 2003, 15:33:16  
Job time : 10.9545 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:22:10 ; Search time 5.18182 seconds  
(without alignments)  
48.025 Million cell updates/sec

Title: US-09-543-188a-4  
Sequence: 1 WLXWIP 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Swissprot\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	88.6	662	1	LOX2_RAT
2	39	88.6	662	1	LOX2_MOUSE
3	37	84.1	340	1	CATS_MOUSE
4	37	84.1	415	1	PTN2_HUMAN
5	37	84.1	2701	1	IP3S_HUMAN
6	37	84.1	2701	1	IP3S_RAT
7	36	81.8	456	1	MGAL_YEAST
8	36	81.8	564	1	YQ2L_CAEEL
9	35	79.5	300	1	FLIR_SALTY
10	35	79.5	305	1	VO41_HSV1
11	35	79.5	360	1	AOYL_YEAST
12	35	79.5	410	1	MRAY_HAFIN
13	35	79.5	467	1	CSG2_YEAST
14	35	79.5	467	1	MMOP_ECOLI
15	35	79.5	467	1	ROCE_BACSU
16	35	79.5	528	1	COX1_TRIRU
17	35	79.5	538	1	COX1_RHILE
18	35	79.5	541	1	COX1_BRAJA
19	35	79.5	541	1	COX1_TODAN
20	34	77.3	40	1	YB5L_PLEBO
21	34	77.3	187	1	Y618_ARCFU
22	34	77.3	268	1	TIFA_ARATH
23	34	77.3	322	1	AOP4_MOUSE
24	34	77.3	323	1	AOP4_BOVIN
25	34	77.3	323	1	AOP4_HUMAN
26	34	77.3	323	1	AOP4_RAT
27	34	77.3	433	1	YBL4_YEAST
28	34	77.3	439	1	VAL1_YEAST
29	34	77.3	627	1	TESK_MOUSE
30	34	77.3	821	1	HAH1_DEBOC
31	34	77.3	2164	1	CCAA_MOUSE
32	34	77.3	2212	1	CCAA_RAT
33	34	77.3	2216	1	YCF2_EPIVI

34	34	77.3	2222	1	CCAE_RAT	Q07652	rattus norv
35	34	77.3	2223	1	CCAE_DISOM	P56699	discopyle o
36	34	77.3	2259	1	CCAE_RABIT	O02343	oryctolagus
37	34	77.3	2272	1	CCAE_MOUSE	O61290	mus musculu
38	34	77.3	2312	1	CCAE_HUMAN	Q15878	homo sapien
39	34	77.3	2326	1	CCAB_DISOM	P56998	discopyle o
40	34	77.3	2327	1	CCAB_MOUSE	O05017	mus musculu
41	34	77.3	2336	1	CCAB_RAT	O02294	rattus norv
42	34	77.3	2339	1	CCAB_HUMAN	O00975	homo sapien
43	34	77.3	2339	1	CCAB_RABIT	O05152	oryctolagus
44	34	77.3	2424	1	CCAA_RABIT	P27884	oryctolagus
45	34	77.3	2505	1	CCAA_HUMAN	O00555	homo sapien

## ALIGNMENTS

RESULT 1	ID	LOX2_RAT	STANDARD	PRT	662 AA
AC	002759	LOX2_RAT			
DT	01-JUL-1993	(Rel. 26, Created)			
DT	01-JUL-1993	(Rel. 26, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Arachidonate 12-lipoxygenase (EC 1.13.11.31) (12-LOX).				
GN	ALOX12.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RP	[1]				
RC	SEQUENCE FROM N.A.				
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;				
RX	MEDLINE=93185682; PubMed8444196;				
RA	Watanabe T., Medina J.F., Haegstroem J.Z., Raadmark O.P.,				
RA	Samuelsson B.;				
RT	"Molecular cloning of a 12-lipoxygenase cDNA from rat brain.";				
RL	Eur. J. Biochem. 212:605-612(1993).				
RP	[2]				
RC	SEQUENCE FROM N.A.				
RC	TISSUE=Pinel gland;				
RA	MEDLINE=94162305; PubMed-8117750;				
RA	Hada T., Hagiya H., Suzuki H., Arikawa T., Nakamura M.,				
RA	Matsuda S., Yoshimoto T., Yamamoto S., Aekawa T., Morita Y.,				
RT	Ishimura K., Kim H.Y.;				
RT	"Arachidonate 12-lipoxygenase of rat pineal glands: catalytic				
RT	properties and primary structure deduced from its cDNA.";				
RL	Biochim. Biophys. Acta 1211:221-228(1994).				
CC	-1- FUNCTION: OXYGENASE AND 14,15-LEUKOTRIENE A4 SYNTHASE ACTIVITY.				
CC	-1- CATALYTIC ACTIVITY: Arachidonate + O(2) = (5Z,8Z,10E,14Z)-(12S)-				
CC	12-hydroperoxyicoso-5,8,10,14-tetraenoate.				
CC	-1- CORFACTOR: IRON.				
CC	-1- PATHWAY: Leukotrienes biosynthesis.				
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.				
CC	-1- SIMILARITY: BELONGS TO THE LIPOXYGENASE FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
DR	EMBL; L06040; AAA1532.1; -				
DR	EMBL; S69363; AAB30132.1; -				
DR	PIR; S30051; S30051.				
DR	HSSP; P12530; ILOX.				
DR	InterPro; IPR000907; Lipoxygenase.				
DR	InterPro; IPR001024; Lipoxygenase_LH2.				
DR	Pfam; PF00305; Lipoxygenase; 1.				
DR	Pfam; PF01477; PLAT; 1.				
DR	PRINTS; PR00087; LIPOXYGENASE.				

```

DR SMART; SM00308; LH2; 1.
DR PROSITE; PS00081; LIPOXYGENASE_2; 1.
DR PROSITE; PS00711; LIPOXYGENASE_1; 1.
KW Oxidoreductase; Dioxxygenase; Iron; Leukotriene biosynthesis.
FT INIT_MET 0
FT METAL 360 360 IRON (BY SIMILARITY).
FT METAL 365 365 IRON (BY SIMILARITY).
FT METAL 540 540 IRON (BY SIMILARITY).
FT METAL 662 662 IRON (BY SIMILARITY).
FT CONFLICT 54 54 E -> G (IN REF. 2).
FT CONFLICT 370 370 L -> V (IN REF. 2).
SQ SEQUENCE 662 AA; 75260 MW; C0187FMAA27FE8F38 CRC64;

Query Match
Best Local Similarity 88.6%; Score 39; DB 1; Length 662;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLWIP 6
550 WEYWP 555

RESULT 2
LOXL_MOUSE
ID LOXL_MOUSE STANDARD; PRT; 662 AA.
AC P39654;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Arachidonate 12-11poxxygenase, leukocyte-type (EC 1.13.11.31) (12-LOX).
GN ALOX12L OR ALOX15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6, and ICR; TISSUE=Spleen;
RX MEDLINE=94245713; PubMed=8188678;
RA Chen X.-S., Kurte U., Jenkins N.A., Copeland N.G., Funk C.D.;
RT "cDNA cloning, expression, mutagenesis of C-terminal isoleucine,
RT genomic structure, and chromosomal localizations of murine
RT 12-11poxxygenase.";
RL J. Biol. Chem. 269:13979-13987(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Macrophage;
RX MEDLINE=95110857; PubMed=7811740;
RA Freire-Moar J., Alavi-Nassab A., Ng M., Mulkins M., Sigal E.;
RT "Cloning and characterization of a murine macrophage lipoxxygenase.";
RL Biochim. Biophys. Acta 1254:112-116(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC CONVERTS ARACHIDONIC ACID TO 12(S)-HYDRO(PERO)XYEICOSA-TETRAENOIC
ACID (12-HPETE) AND 15-HYDRO(PERO)XYEICOSA-TETRAENOIC ACID IN A 3:1
RATIO. CONVERTS AS WELL LINOLEIC ACID TO 13-HYDRO-
(PERO)XYCOTADECADENOIC ACID.
CC -1- CATALYTIC ACTIVITY: Arachidonate + O(2) -> (5Z,8Z,10E,14Z)-(12S)-
CC 12-hydroperoxyicoso-5,8,10,14-tetraenoate.
CC -1- COFACTOR: IRON.
CC -1- PATHWAY: Leukotrienes biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: FOUND IN PITUITARY AND PINEAL GLANDS AS WELL
CC AS LEUKOCYTES, KIDNEY, AORTA, AND SMALL INTESTINE.
CC -1- SIMILARITY: BELONGS TO THE LIPOXYGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; U04331; AAA20658.1; -.
DR EMBL; L34570; AAA64930.1; -.
DR HSSP; P12530; ILOX.
DR MGD; MGI:87997; Alox15.
DR InterPro; IPR000907; Lipoxxygenase.
DR InterPro; IPR001024; Lipoxxygenase_LH2.
DR Pfam; PF00305; lipoxxygenase; 1.
DR Pfam; PF01477; PLAT; 1.
DR PRINTS; PR00087; LIPOXYGENASE.
DR SMART; SM00308; LH2; 1.
DR PROSITE; PS00081; LIPOXYGENASE_2; 1.
DR PROSITE; PS00711; LIPOXYGENASE_1; 1.
KW Oxidoreductase; Dioxxygenase; Iron; Leukotriene biosynthesis.
FT INIT_MET 0
FT METAL 360 360 IRON (BY SIMILARITY).
FT METAL 365 365 IRON (BY SIMILARITY).
FT METAL 540 540 IRON (BY SIMILARITY).
FT METAL 662 662 IRON (BY SIMILARITY).
FT CONFLICT 36 36 K -> N (IN REF. 2).
FT CONFLICT 118 118 E -> Q (IN REF. 2).
FT CONFLICT 396 396 T -> N (IN REF. 2).
SQ SEQUENCE 662 AA; 75313 MW; 5C94965B30767C2C CRC64;

Query Match
Best Local Similarity 88.6%; Score 39; DB 1; Length 662;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLWIP 6
550 WEYWP 555

RESULT 3
CATS_MOUSE
ID CATS_MOUSE STANDARD; PRT; 340 AA.
AC 070370; O54973;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cathepsin S precursor (EC 3.4.22.27).
GN CTSS OR CATS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/c, and 129/Sv; TISSUE=Brain;
RA Doh-ura K.;
RT Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Rommelskirch W.;
RT submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 144-306 FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=C57BL/6; TISSUE=Cartilage;
RX MEDLINE=99326135; PubMed=10395917;
RA Soederstrom M., Salminen H., Glunoff V., Kirschke H., Aro H.,
RA Vuorio E.;
RT "Cathepsin expression during skeletal development.";
RL Biochim. Biophys. Acta 1446:35-46(1999).
RN [4]
RP SEQUENCE OF 296-340 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=9818482; PubMed=9516475;
RA Dandoy-Dron F., Guillo F., Benboudjema L., Deslys J.-P., Lasmesas C.,
RA Dormont D., Tovey M.G., Dron M.;
RT "Gene expression in scrapie. Cloning of a new scrapie-responsive gene
RT and the identification of increased levels of seven other mRNA
RT transcripts.";
RL J. Biol. Chem. 273:7691-7697(1998).

```

CC -1- FUNCTION: THIOL PROTEASE. THE BOND-SPECIFICITY OF THIS PROTEINASE  
 CC IS IN PART SIMILAR TO THE SPECIFICITIES OF CATHEPSIN L AND  
 CC CATHEPSIN N.  
 CC -1- CATALYTIC ACTIVITY: Similar to cathepsin L, but with much less  
 CC activity on 2-Phe-Arg-1-NHec, and more activity on the 2-Val-Val-  
 CC Arg-1-Xaa compound.  
 CC -1- SUBCELLULAR LOCATION: Lysosomal.  
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHEST EXPRESSION FOUND  
 CC IN NON-SKELETAL TISSUES. RELATIVELY HIGH LEVELS FOUND IN SKELETAL  
 CC TISSUES.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF051732; AAC05781.1; -;  
 CC EMBL: AF051727; AAC05781.1; JOINED.  
 CC EMBL: AF051728; AAC05781.1; JOINED.  
 CC EMBL: AF051729; AAC05781.1; JOINED.  
 CC EMBL: AF051726; AAC05781.1; JOINED.  
 CC EMBL: AF051730; AAC05781.1; JOINED.  
 CC EMBL: AF051731; AAC05781.1; JOINED.  
 CC EMBL: AF038546; AAB94925.1; -;  
 CC EMBL: AJ002386; CAA05360.1; -;  
 CC EMBL: Y18466; CAA77184.1; -;  
 CC EMBL: AJ23208; CAA11182.1; -;  
 CC HSSP: P25774; 1BXF.  
 CC DR MEROPS: C01.034; -;  
 CC DR MGD: MG1.107341; Ctss.  
 CC DR InterPro: IPR000668; Peptidase\_C1.  
 CC DR InterPro: IPR000169; SHprol\_acstite.  
 CC DR Pfam: PF00112; Peptidase\_C1; 1.  
 CC DR PRINTS: PR00705; PAPA1N.  
 CC DR Prodom: PD000158; Peptidase\_C1; 1.  
 CC DR PROSITE: PS00139; THIOL\_PROTEASE\_CYS; 1.  
 CC DR PROSITE: PS00639; THIOL\_PROTEASE\_HIS; 1.  
 CC DR PROSITE: PS00640; THIOL\_PROTEASE\_ASN; 1.  
 CC KW Hydrolase; Thiol protease; Lysosome; Zymogen; Signal.  
 CC FT SIGNAL 1 17  
 CC FT PROPEP 18 112  
 CC FT CHAIN 113 340  
 CC FT ACT\_SITE 147 147  
 CC FT ACT\_SITE 287 287  
 CC FT ACT\_SITE 307 307  
 CC FT DISULFID 134 233  
 CC FT DISULFID 144 189  
 CC FT DISULFID 178 232  
 CC FT DISULFID 281 329  
 CC FT CARBOHYD 120 120  
 CC FT VARIANT 218 218  
 CC FT CONFLICT 1 29  
 CC FT CONFLICT 34 34  
 CC FT CONFLICT 97 97  
 CC FT CONFLICT 106 106  
 CC FT CONFLICT 146 146  
 CC SQ SEQUENCE 340 AA; 38438 MW; 068B611262E0C0E CRC64;  
 CC  
 CC Query Match Best Local Similarity 66.7%; Score 37; DB 1; Length 340;  
 CC Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 1 WLYWIP 6  
 CC 11:1:1  
 CC Db 11 WLFWMP 16  
 CC  
 CC RESULT 4

PTN2\_HUMAN  
 ID PTN2\_HUMAN STANDARD; PRT; 415 AA.  
 AC P17706; O96HR2;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Protein-tyrosine phosphatase, non-receptor type 2 (BC 3.1.3.48) (T-  
 DE cell protein-tyrosine phosphatase) (TCPTP).  
 GN PTNP2 OR PTPP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SOURCE FROM N.A.  
 RC TISSUE=T-cell;  
 RA MEDLINE=9315776; PubMed=2546150;  
 RX Cool D., Tonks N., Chardonneau H., Walsh K., Fischer E.H., Krebs E.G.;  
 RT "CDNA isolated from a human T-cell library encodes a member of the  
 RT protein-tyrosine-phosphatase family";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:5257-5261(1989).  
 RP [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92115688; PubMed=1731319;  
 RA Mosinger B. Jr., Tillmann U., Westphal H., Tremblay M.L.;  
 RT "Cloning and characterization of a mouse cDNA encoding a cytoplasmic  
 RT protein-tyrosine-phosphatase";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:499-503(1992).  
 RP [3]  
 RP SEQUENCE FROM N.A. (ISOFORM PTPB).  
 RC TISSUE=Eye;  
 RA Strausberg R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein  
 CC tyrosine + phosphate.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: PTPA AND PTPB (SHOWN HERE); ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: PTPA ISOFORM IS PROBABLY THE MAJOR PTP  
 CC EXPRESSED IN HUMAN TISSUES. PTPB ISOFORM WAS FOUND IN T-CELLS AND  
 CC IN PLACENTA.  
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-  
 CC TYROSINE PHOSPHATASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M25393; AAA65997.1; -;  
 CC EMBL: M81478; -; NOT\_ANNOTATED\_CDS.  
 CC EMBL: BC008244; AAH08244.1; -;  
 CC PIR: A33899; A33899.  
 CC HSSP: P18031; 1PTY.  
 CC DR Genew; HGNC:9650; PTPN2.  
 CC DR MIM: 176887; -;  
 CC DR InterPro: IPR000387; Tyr\_P.  
 CC DR InterPro: IPR000242; Tyr\_PP.  
 CC DR Pfam: PF00102; Y-phosphatase; 1.  
 CC DR PRINTS: PR00700; PRTYPPHASE.  
 CC DR SMART: SM00194; PTPc; 1.  
 CC DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 1.  
 CC DR PROSITE: PS00056; TYR\_PHOSPHATASE\_2; 1.  
 CC DR PROSITE: PS00055; TYR\_PHOSPHATASE\_PTP; 1.  
 CC KW Hydrolase; T-cell; Alternative splicing.  
 CC FT DOMAIN 42 286  
 CC FT ACT\_SITE 216 216  
 CC FT VARSPLIC 382 415  
 CC FT  
 CC SQ SEQUENCE 415 AA; 48528 MW; 0599694A4A4F058E68 CRC64;  
 CC  
 CC TDR (IN ISOFORM PTPA).  
 CC WLYWIPILTKMGFMSYLVGAFVGRILFQGNAL -> PRL  
 CC FT

Query Match 84.1%; Score 37; DB 1; Length 415;  
Best Local Similarity 83.3%; Pred. No. 28;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLWMP 6  
11111  
Db 382 WLWMP 387

## RESULT 5

IP3S\_HUMAN STANDARD; PRT; 2701 AA.  
ID IP3S\_HUMAN  
AC 014571; 094773;  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Inositol 1,4,5-trisphosphate receptor type 2 (Type 2 inositol 1,4,5-  
DE trisphosphate receptor) (Type 2 InsP3 receptor) (IP3 receptor isoform  
2) (InsP3R2).  
ITPR2.  
Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RP SEQUENCE FROM N.A. (LONG ISOFORM).  
RX MEDLINE=94363219; PubMed=8081773;  
RA Yamamoto-Hino M., Sugiyama T., Hikita K., Mattei M.-G.,  
RA Hasegawa K., Sekine S., Sakurada K., Miyawaki A., Furutachi T.,  
RA Hasegawa M., Mikoshiba K.;  
RT "Cloning and characterization of human type 2 and type 3 inositol  
RT 1,4,5-trisphosphate receptors.";  
RL Recept. Channels 2:9-22(1994).  
[2]  
RN SEQUENCE FROM N.A. (SHORT ISOFORM).  
RP TISSUE=Heart;  
RX MEDLINE=98399819; PubMed=9729462;  
RA Futatsugi A., Kuwajima G., Mikoshiba K.;  
RT "Muscle-specific mRNA isoform encodes a protein composed mainly of the  
RT N-terminal 175 residues of type 2 Ins(1,4,5)P3 receptor.";  
RL Biochem. J. 334:559-563(1998).  
CC -1- FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND  
CC MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.  
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic  
CC reticulum.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM/TIPR; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: THE SHORT ISOFORM IS FOUND IN SKELETAL MUSCLE  
CC AND HEART.  
CC -1- DOMAIN: THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL  
CC EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-  
CC BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE  
CC PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.  
CC -1- PTM: PHOSPHORYLATED ON TYROSINE RESIDUES (BY SIMILARITY).  
CC -1- MISCELLANEOUS: CALCIUM APPEARS TO INHIBIT LIGAND BINDING TO THE  
CC RECEPTOR, MOST PROBABLY BY INTERACTING WITH A DISTINCT  
CC CALCIUM-BINDING PROTEIN WHICH THEN INHIBITS THE RECEPTOR.  
CC -1- SIMILARITY: BELONGS TO THE INSP3 RECEPTOR FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: D26350; BAA05384.1;  
CC EMBL: AB012610; BAA33961.1;  
CC DR Genew; HGNC:6181; ITPR2.  
CC MIM: 600144; .

DR InterPro: IPR000699; Ca-rel\_channel.  
DR InterPro: IPR001682; Ca/Na\_pore.  
DR InterPro: IPR000493; InsP3\_receptor.  
DR InterPro: IPR000636; M+channel\_n19.  
DR InterPro: IPR003608; MIR.  
DR Pfam: PF00520; Ion\_trans\_1.  
DR Pfam: PF01365; RyDR-IPPR; 2.  
DR Pfam: PF02815; MIR; 4.  
DR PRINTS: PR00779; INSP3RECEPTR.  
DR SMART: SM00472; MIR; 4.  
KW Receptor; Transmembrane; Glycoprotein; Phosphorylation;  
KW Endoplasmic reticulum; Ionic channel; Ion transport; Calcium channel;  
KW Alternative splicing.  
FT DOMAIN 1 2227 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 2228 2248 POTENTIAL.  
FT DOMAIN 2249 2260 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 2261 2281 POTENTIAL.  
FT DOMAIN 2282 2307 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 2308 2328 POTENTIAL.  
FT DOMAIN 2329 2351 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 2352 2372 POTENTIAL.  
FT DOMAIN 2373 2394 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 2395 2415 POTENTIAL.  
FT DOMAIN 2416 2521 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 2522 2542 POTENTIAL.  
FT DOMAIN 2543 2701 CYTOPLASMIC (POTENTIAL).  
FT MOD\_RES 2607 2607 PHOSPHORYLATION (POTENTIAL).  
FT VARSPIC 176 181 IYVGRK->DASFRI (IN SHORT ISOFORM).  
FT VARSPIC 182 2701 MISSING (IN SHORT ISOFORM).  
SQ SEQUENCE 2701 AA; 308074 MW; EB5C7DDCD17F7A4 CRC64;

Query Match 84.1%; Score 37; DB 1; Length 2701;  
Best Local Similarity 100.0%; Pred. No. 1,6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWMI 5  
11111  
Db 694 WLWMI 698

## RESULT 6

IP3S\_RAT STANDARD; PRT; 2701 AA.  
ID IP3S\_RAT  
AC P29995; 099P56;  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Inositol 1,4,5-trisphosphate receptor type 2 (Type 2 inositol 1,4,5-  
DE trisphosphate receptor) (Type 2 InsP3 receptor) (IP3 receptor isoform  
2) (InsP3R2).  
GN ITPR2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
[1]  
RP SEQUENCE FROM N.A. AND VARIANTS.  
RC TISSUE=Brain;  
RX MEDLINE=92007769; PubMed=1655411;  
RA Suedhof T.C., Newton C.A., Archer B.T. III, Usinkayov Y.A.,  
RA Mignery G.A.;  
RT "Structure of a novel InsP3 receptor.";  
RL EMBO J. 10:3199-3206(1991).  
[2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Sprague-Dawley;  
RC Magnino F., Dufour J.-F.;  
RT "New rat IP3R isoform 2 sequence.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND  
CC MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.  
CC -1- SUBUNIT: HOMOTETRAMER.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic



```
CC reticulum.
CC -1- DOMAIN: THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL
CC EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-
CC BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE
CC PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.
CC -1- PHM: PHOSPHORYLATED ON TYROSINE RESIDUES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSP3 RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X61677; CAA43852.1; -.
CC EMBL: AF329470; AAK11622.1; -.
CC PIR: S17796; S17796.
CC InterPro: IPR000699; Ca-rel_channel.
CC InterPro: IPR001682; Ca/Na_pore.
CC InterPro: IPR000493; INSP3_receptor.
CC InterPro: IPR000636; W-channel_n19.
CC InterPro: PF00520; Ion_trans; 1.
CC Pfam: PF01365; RYDR_ITPR; 2.
CC Pfam: PF02815; MIR; 4.
CC PRINTS: PR00779; INSP3RECEPTR.
CC SMART: SM00472; MIR; 4.
CC Receptor; Transmembrane; Phosphorylation; Endoplasmic reticulum;
CC Ionic channel; Ion transport; Calcium channel; Polymorphism.
CC DOMAIN 1 2227
CC TRANSMM 2228 2248 POTENTIAL.
CC TRANSMM 2249 2260 EXTRACELLULAR (POTENTIAL).
CC TRANSMM 2261 2281 POTENTIAL.
CC TRANSMM 2282 2284 CYTOPLASMIC (POTENTIAL).
CC TRANSMM 2285 2305 POTENTIAL.
CC TRANSMM 2306 2307 EXTRACELLULAR (POTENTIAL).
CC TRANSMM 2308 2328 POTENTIAL.
CC TRANSMM 2329 2351 CYTOPLASMIC (POTENTIAL).
CC TRANSMM 2352 2372 POTENTIAL.
CC TRANSMM 2373 2394 EXTRACELLULAR (POTENTIAL).
CC TRANSMM 2395 2415 POTENTIAL.
CC TRANSMM 2416 2521 CYTOPLASMIC (POTENTIAL).
CC TRANSMM 2522 2542 EXTRACELLULAR (POTENTIAL).
CC TRANSMM 2543 2701 POTENTIAL.
CC TRANSMM 2607 2607 EXTRACELLULAR (POTENTIAL).
CC TRANSMM 2607 2607 PHOSPHORYLATION (POTENTIAL).
CC TRANSMM 689 689 D -> H.
CC TRANSMM 1013 1013 G -> C.
CC TRANSMM 1256 1256 L -> P.
CC TRANSMM 2384 2384 V -> I.
CC TRANSMM 2694 2694 E -> V.
CC TRANSMM 119 119 K -> N (IN REF. 2).
CC TRANSMM 344 344 H -> R (IN REF. 2).
CC TRANSMM 943 943 W -> V (IN REF. 2).
CC TRANSMM 1692 1692 S -> G (IN REF. 2).
CC TRANSMM 2556 2556 K -> E (IN REF. 2).
CC TRANSMM 2701 AA; 307054 MW; 42BF7F1024335584 CRC64;
CC SEQUENCE
CC
CC Query Match 84.1%; Score 37; DB 1; Length 2701;
CC Best Local Similarity 100.0%; Pred. No. 1.6e+02;
CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE MGAI protein.
CN MGAI OR YGR249W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DD416-1-4;
RL Wada H.;
RM Submitted (Apr-1994) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=97279234; PubMed=9133742;
RA Feroli F., Carignani G., Pavanello A., Guerreiro P., Azevedo D.,
RA Rodrigues-Pousada C., Melchiorretto P., Panzeri L.,
RA Agostoni Carbone M.L.;
RT "Analysis of a 17.9 kb region from Saccharomyces cerevisiae
RT chromosome VII reveals the presence of eight open reading frames,
RT including BRF1 (YFI1B70) and GCN5 genes."
RL Yeast 13:373-377(1997).
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE HSF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D29626; BAA06105.1; -.
CC EMBL: Y07703; CAA68970.1; -.
CC EMBL: Z73034; CAA97278.1; -.
CC HSP: P22121; 2HTS.
CC TRANSFAC: T03496; -.
CC SGD: S0003481; MGAI.
CC InterPro: IPR002332; HSF_DNA_bind.
CC InterPro: IPR002341; HSF_ETS.
CC Pfam: PF00447; HSF_DNA_bind; 1.
CC PRINTS: PR001788; HSPDOMAIN.
CC PRODOM: PD001788; HSF_DNA_bind; 1.
CC SMART: SM00415; HSF; 1.
CC DR PROSITE: PS00434; HSF_DOMAIN; 1.
CC KW Nuclear protein; DNA-binding.
CC FT DNA_BIND 3 118 BY SIMILARITY.
CC FT DOMAIN 182 185 POLY-GLN.
CC FT DOMAIN 283 290 POLY-GLN.
CC SEQUENCE 456 AA; 50735 MW; 48A21F0D89749BA6 CRC64;
CC
CC Query Match 81.8%; Score 36; DB 1; Length 456;
CC Best Local Similarity 66.7%; Pred. No. 44;
CC Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 WLWWT 5
DB 694 WLWWT 698

RESULT 7
MGAI_YEAST STANDARD: PRT; 456 AA.
AC P53050.
DT 01-OCT-1996 (Rel. 34, Created)
```

```
YQ21_CAEEL
ID YQ21_CAEEL STANDARD: PRT; 564 AA.
AC 009231;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 63.8 kDa protein C09P5.1 in chromosome III.
C09P5.1.
OS Caenorhabditis elegans.
```

CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
CC Rhabditidae; Peloderinae; Caenorhabditis.  
CC NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Bristol N2;  
RA Connell M.;  
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO C.ELEGANS W01C9.3.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC DR EMBL: U28332; AAA64509.1; -  
CC WormPep: C09F5.1; CE01773.  
CC Hypothetical protein.  
CC SEQUENCE 564 AA: 63824 MW: 04A2D5ACFB7C75E1D CRC64;  
-----  
OY 1 WLWTP 6  
DB 539 WWMWP 544  
-----  
Query Match 81.8%; Score 36; DB 1; Length 564;  
Best Local Similarity 66.7%; Pred. No. 54;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
-----  
RESULT 9  
FLIR\_SALTY  
ID FLIR\_SALTY STANDARD: PRT: 264 AA.  
AC P34702;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Flagellar biosynthetic protein flir.  
GN FLIR OR FLAP OR STM1981.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
CC Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-LT2;  
MEDLINE=97464436; PubMed=9324257;  
Ohnishi K., Fan F., Schoenhals G.J., Kihara M., Macnab R.M.;  
"The FljO, FljP, FljQ, and FljR proteins of Salmonella typhimurium:  
RT putative components for flagellar assembly.";  
RL J. Bacteriol. 179:6092-6099(1997).  
-----  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-LT2 / SGSC1412 / ATCC 700720;  
MEDLINE=21534948; PubMed=11677609;  
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,  
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvoney E.,  
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
RA Waterston R., Wilson R.K.;  
"Complete genome sequence of Salmonella enterica serovar Typhimurium  
LT2.";  
RL Nature 413:852-856(2001).  
-----  
CC -1- FUNCTION: ROLE IN FLAGELLAR BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
CC -1- SIMILARITY: BELONGS TO THE FLIR/MOPE/SPAR FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC DR EMBL: L49021; AAB81321.1; -  
CC DR EMBL: AE008787; AAL20893.1; -  
CC DR StyGene: SG10579; flir.  
CC InterPro: IPR002010; Bac\_export\_1.  
CC Pfam: PF01311; Bac\_export\_1; 1.  
CC Flgella; Transmembrane; Inner membrane; Complete proteome.  
FT TRANSMEM 45 65 POTENTIAL.  
FT TRANSMEM 66 86 POTENTIAL.  
FT TRANSMEM 131 151 POTENTIAL.  
FT TRANSMEM 175 195 POTENTIAL.  
FT TRANSMEM 213 233 POTENTIAL.  
FT CONFLICT 24 24 L->R (IN REF. 1).  
SO SEQUENCE 264 AA: 28925 MW: 7AFF46906147A63F CRC64;  
-----  
Query Match 79.5%; Score 35; DB 1; Length 264;  
Best Local Similarity 80.0%; Pred. No. 38;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
-----  
OY 1 WLWTP 5  
DB 9 WML 13  
-----  
RESULT 10  
VG41\_HSV11  
ID VG41\_HSV11 STANDARD: PRT: 300 AA.  
AC Q00017;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 01-DEC-1992 (Rel. 24, Last annotation update)  
DE Hypothetical gene 41 protein.  
GN 41.  
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
CC Ictalurid Herpes-like viruses.  
OX NCBI\_TaxID=10401;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Auburn 1;  
MEDLINE=92087490; PubMed=1727613;  
RA Davison A.J.;  
"Channel catfish virus: a new type of herpesvirus.";  
RL Virology 186:9-14(1992).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC DR EMBL: M75136; AAA8144.1; -  
CC DR PIR: F36790; F36790.  
CC Hypothetical protein.  
SO SEQUENCE 300 AA: 33892 MW: D3179E15F0F3A6D CRC64;  
-----  
Query Match 79.5%; Score 35; DB 1; Length 300;  
Best Local Similarity 66.7%; Pred. No. 43;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
-----  
OY 1 WLWTP 6  
DB 57 WWMWP 62  
-----  
RESULT 11  
AOYL\_YEAST  
ID AOYL\_YEAST STANDARD: PRT: 305 AA.  
-----

AC P53386; 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Aquaporin-like protein AQP1.  
GN AQP1 OR YP192W OR P9677.5.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=97313271; PubMed=9169875;  
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,  
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Bene V.,  
RA Bostein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,  
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,  
RA Dietrich F.S., Dilius H., DiPaolo T., Dubois E., Duesterhoeft A.,  
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,  
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,  
RA Hunke-Smith S., Hyman R., Johnston M., Kalman S., Kaine K.,  
RA Komp C., Kurd O., Lastkari D., Lew H., Lin A., Lin D., Louis E.J.,  
RA Marathe R., Messenguy F., Mewes H.-W., Miliptat S., Moestl D.,  
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,  
RA Petel F.X., Pohl T.M., Purnelle D., Schaefer M., Scharif M.,  
RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,  
RA Uristarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,  
RA Walsh S.V., Walther R., Wang Y., Wedler E., Wedler H., Winnett E.,  
RA Zhong W.W., Zolnier A., Vo D.H., Hanl J.;  
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";  
RA Nature 387:103-105(1997).  
RN [2]  
RP CHARACTERIZATION.  
RX MEDLINE=96438534; PubMed=9765289;  
RA Bonhivers M., Cardrey J.M., Gould S.J., Agre P.;  
RA "Aquaporins in Saccharomyces. Genetic and functional distinctions  
RT between laboratory and wild-type strains.";  
RL J. Biol. Chem. 273:27565-27572(1998).  
CC -1- FUNCTION: WATER-SPECIFIC CHANNEL. SEEMS TO BE NONFUNCTIONAL IN  
CC LABORATORY STRAINS.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO THE MIP/AQUAPORIN FAMILY (NC 1.A.8).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: U25841; AA664621.1; -  
CC HSSP: P29972; IFOY.  
CC SGD: S0006396; AQP1.  
CC InterPro: IPR000425; MIP\_family.  
CC Pfam: PF00230; MIP; 1.  
CC ProDom: PD000295; MIP\_family; 1.  
CC TIGRfams: TIGR00861; MIP; 1.  
CC PROSITE: PS00221; MIP; 1.  
CC Transport: Transmembrane.  
CC TRANSMEM 49 69 POTENTIAL.  
CC TRANSMEM 92 112 POTENTIAL.  
CC TRANSMEM 137 157 POTENTIAL.  
CC TRANSMEM 177 197 POTENTIAL.  
CC TRANSMEM 204 224 POTENTIAL.  
CC TRANSMEM 249 269 POTENTIAL.  
CC VARIANT 121 121 M -> V (IN WILD STRAINS).  
CC VARIANT 255 255 T -> P (IN WILD STRAINS).  
CC SEQUENCE 305 AA; 32712 MW; 2C02491CF93AC7F CR64;  
Query Match 79.5%; Score 35; DB 1; Length 305;  
Best Local Similarity 80.0%; Pred. No. 44;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 WITWI 5  
Db 249 WITWI 253  
RESULT 12  
MRAY\_HAEIN STANDARD: PRT: 360 AA.  
ID MRAY\_HAEIN  
AC P45062;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Phospho-N-acetylmutaroyl-pentapeptidyl transferase (EC 2.7.8.13) (UDP-  
DE MurNAc-pentapeptide phosphotransferase).  
GN MRAY OR H1135.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Rd / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischnann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Metrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Uitterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Georgagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae  
RT Rd.";  
RL Science 269:496-512(1995).  
CC -1- FUNCTION: FIRST STEP OF THE LIPID CYCLE REACTIONS IN THE  
CC BIOSYNTHESIS OF THE CELL WALL. PEPTIDOGLYCAN (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: UDP-N-acetylmuramoyl-L-alanyl-D-glutamyl-L-  
CC lysyl-D-alanyl-D-alanine + undecaprenyl phosphate = UDP + N-  
CC acetylmuramoyl-L-alanyl-D-glutamyl-L-lysyl-D-alanyl-D-alanine-  
CC diposphundecaprenol.  
CC -1- PATHWAY: Peptidoglycan biosynthesis.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).  
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 4. MRAY  
CC SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: U32793; AAC22790.1; -  
CC TIGR: H1135;  
CC InterPro: IPR000715; Glycos\_transf\_4.  
CC InterPro: IPR003524; PNAcp\_transf.  
CC Pfam: PF00953; Glycos\_transf\_4; 1.  
CC TIGRfams: TIGR00445; mray; 1.  
CC PROSITE: PS01347; MRAY\_1; 1.  
CC PROSITE: PS01348; MRAY\_2; 1.  
CC Peptidoglycan synthesis; Cell division; Transferase; Transmembrane;  
CC Complete proteome.  
CC TRANSMEM 21 41 POTENTIAL.  
CC TRANSMEM 73 93 POTENTIAL.  
CC TRANSMEM 94 114 POTENTIAL.  
CC TRANSMEM 132 152 POTENTIAL.  
CC TRANSMEM 168 188 POTENTIAL.  
CC TRANSMEM 200 220 POTENTIAL.  
CC TRANSMEM 239 259 POTENTIAL.

FT	TRANSMEM	263	283	POTENTIAL.
FT	TRANSMEM	288	308	POTENTIAL.
FT	TRANSMEM	338	358	POTENTIAL.
SO	SEQUENCE	360 AA;	40315 MW;	C10D066ADCA0508D CRC64:

Query Match 79.5%; Score 35; DB 1; Length 360;  
Best Local Similarity 80.0%; Pred. No. 51;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WLYWT 5  
11111:  
Db 147 WLYWL 151

RESULT 13  
CSG2\_YEAST  
ID CSG2\_YEAST STANDARD: PRT: 410 AA.  
AC P35206:  
DT 01-FEB-1994 (Rel. 28, Created)  
01-FEB-1994 (Rel. 28, Last sequence update)  
16-OCT-2001 (Rel. 40, Last annotation update)  
CSG2 protein precursor.  
GN CSG2 OR CLS2 OR YBR036C OR YBR0404.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94171742; PubMed=8125941;  
RA Beeler T.J., Gable K., Zhao C., Dunn T.;  
RT "A novel protein, CSG2p, is required for Ca2+ regulation in  
RL Saccharomyces cerevisiae.";  
RN J. Biol. Chem. 269:7279-7284(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=5288C;  
RA Andre B., Cziepluch C., Hein C., Jauniaux J.C., Urrestarazu A.,  
RL Vissers S.;  
RN Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95157516; PubMed=7854312;  
RA Takita Y., Ohya Y., Anraku Y.;  
RT "The CLS2 gene encodes a protein with multiple membrane-spanning  
RN domains that is important Ca2+ tolerance in yeast.";  
RL Mol. Gen. Genet. 246:269-281(1995).  
RN [4]  
RP SEQUENCE OF 116-410 FROM N.A.  
RX STRAIN=5288C;  
RA MEDLINE=94378725; PubMed=8091864;  
RA Smits P.H.M., de Haan M., Maat C., Grivell L.A.;  
RT "The complete sequence of a 33 kb fragment on the right arm of  
RN chromosome II from Saccharomyces cerevisiae reveals 16 open reading  
RT frames, including ten new open reading frames, five previously  
RN identified genes and a homologue of the SCO1 gene.";  
RL Yeast 10:S75-S80(1994).  
RN [5]  
RP SEQUENCE OF 1-92 FROM N.A.  
RX MEDLINE=89281488; PubMed=2543907;  
RA Schulze M., Koedel G.;  
RT "Accumulation of the cytochrome c oxidase subunits I and II in yeast  
RN requires a mitochondrial membrane-associated protein, encoded by the  
RL nuclear SCO1 gene.";  
RN Mol. Gen. Genet. 216:37-43(1989).  
CC -1- FUNCTION: REQUIRED FOR CALCIUM REGULATION. MAY REGULATE CALCIUM  
CC ACCUMULATION BY A NON-VACUOLE ORGANELLE.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

Query Match	Best Local Similarity	Score 35:	DB 1:	Length 410:
Matches 4:	Conservative	1:	Mismatches 1:	Indels 0:
Gaps 0:				
Oy 1 WLMWP 6	1111			
Db 65 WLMWP 70				
RESULT 14				
MMUP_ECOLI				
ID MMUP_ECOLI	STANDARD:	PRT:	467 AA.	
AC 047689:				
DT 01-NOV-1997 (Rel. 35, Created)				
DT 16-OCT-2001 (Rel. 40, Last sequence update)				
DT 16-OCT-2001 (Rel. 40, Last annotation update)				
DE Probable S-methylmethionine permease.				
GN MMUP OR B0260.				
OS Escherichia coli.				
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				
OC Escherichia.				
OX NCBI_TaxID=562:				
LN [1]				
RP SEQUENCE FROM N.A.				
RC STRAIN-K12 / MG1655.				
RX MEDLINE=9742617; PubMed=9278503;				
RX Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,				
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,				
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,				
RA Mau B., Shao Y.;				
RT "The complete genome sequence of Escherichia coli K-12.";				
RL Science 277:1453-1474(1997).				
RN [2]				
RP SEQUENCE FROM N.A.				
RC STRAIN-K12 / W3110;				
RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,				
RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,				

RA Mizuno T., Makino K., Nakata A., Yura T., Sampei G., Mizobuchi K.;  
 RT "Systematic sequencing of the *Escherichia coli* genome: analysis of the  
 4.0 - 6.0 min (189,987 - 281,416bp) region."  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,  
 RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kundi O.,  
 RA Laskhari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,  
 RA Davis R.W.;  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP POSSIBLE FUNCTION.  
 RX MEDLINE=99102233; PubMed=9882684;  
 RA Thambichler M., Neuhiel B., Boeck A.;  
 RL "S-methylmethionine metabolism in *Escherichia coli*."  
 CC J. Bacteriol. 181:662-665(1999).  
 CC -1- FUNCTION: TRANSPORTER FOR THE INTAKE OF S-METHYLMETHIONINE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE000134; AAC73363.1; ALT\_INIT.  
 DR EMBL: D83536; BAA77928.1; ALT\_INIT.  
 DR EMBL: U70214; AAB08681.1; ALT\_INIT.  
 DR EcoGene: EG33342; mmup.  
 DR InterPro: IPR002293; AA/rel\_pmease1.  
 DR InterPro: IPR004840; AAC\_permease.  
 DR InterPro: IPR004841; Permease.  
 DR Pfam: PF00324; aa\_permeases; 1.  
 DR PROSITE: PS00218; AMINO ACID PERMEASE; 1.  
 DR TRANSPORT; Transmembrane; Inner membrane; Complete proteome.  
 FT TRANSMEM 20 40  
 FT TRANSMEM 43 63 POTENTIAL.  
 FT TRANSMEM 96 116 POTENTIAL.  
 FT TRANSMEM 125 145 POTENTIAL.  
 FT TRANSMEM 161 181 POTENTIAL.  
 FT TRANSMEM 206 226 POTENTIAL.  
 FT TRANSMEM 245 265 POTENTIAL.  
 FT TRANSMEM 282 302 POTENTIAL.  
 FT TRANSMEM 337 357 POTENTIAL.  
 FT TRANSMEM 361 381 POTENTIAL.  
 FT TRANSMEM 408 428 POTENTIAL.  
 FT TRANSMEM 434 454 POTENTIAL.  
 SQ SEQUENCE 467 AA; 50525 MW; E7ADB9A3B03C10F8 CRC64;  
 Query Match 79.5%; Score 35; DB 1; Length 167;  
 Best Local Similarity 80.0%; Pred. No. 65;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 WLWYL 5  
 DB 97 WLWYL 101  
 RESULT 15  
 ID ROCE\_BACSU STANDARD; PRT; 467 AA.  
 AC P39137;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Amino-acid permease roce.  
 GN ROCE.  
 OS *Bacillus subtilis*.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=95311309; PubMed=7540694;  
 RA Gardan R., Rapoport G., Debarouille M.;  
 RT "Expression of the roceDEF operon involved in arginine catabolism in  
 RT *Bacillus subtilis*."  
 RL J. Mol. Biol. 249:843-856(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Kasahara Y., Nakai S., Yoshikawa H., Ogasawara N.;  
 RT "36kb sequence between gntZ and trnY of *B. subtilis* genome."  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertolo M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschel C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Ertlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,  
 RA Gilm S.Y., Glaser P., Gotreau A., Gollightly E.J., Grandi G.,  
 RA Guilseppe G., Guy B.J., Haga K., Haeck J., Hartwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kunano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Meuel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetle D., Porwollik S., Prescott A.M.,  
 RA Pirescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield S., Soldo B.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terstra P., Togroni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vanier F., Vassarotti A.,  
 RA Viari A., Wambuit R., Wedler E., Wedler H., Weitzenecker T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zumbstein A.;  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*."  
 RL Nature 390:249-256(1997).  
 CC -1- FUNCTION: PUTATIVE TRANSPORT PROTEIN INVOLVED IN ARGININE  
 CC DEGRADATIVE PATHWAY. PROBABLY TRANSPORTS ARGININE OR ORNITHINE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- INDUCTION: EXPRESSION IS SIGMA I DEPENDENT, INDUCED BY ARGININE,  
 CC ORNITHINE OR PROLINE.  
 CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X81802; CA57399.1;  
 DR EMBL: D78193; BAA11292.1;  
 DR EMBL: Z99124; CAB16070.1;  
 DR SdbList: BG10933; roce.  
 DR InterPro: IPR002293; AA/rel\_pmease1.  
 DR InterPro: IPR004840; AAC\_permease.  
 DR InterPro: IPR004841; Permease.

DR Pfam: PF00324; aa\_permeases; 1.  
 DR PROSITE: PS00218; AMINO-ACID\_PERMEASE\_1; 1.  
 KW Transport; Amino-acid transport; Transmembrane; Complete proteome.  
 FT TRANSMEM 21 41 POTENTIAL.  
 FT TRANSMEM 47 67 POTENTIAL.  
 FT TRANSMEM 87 107 POTENTIAL.  
 FT TRANSMEM 122 142 POTENTIAL.  
 FT TRANSMEM 162 182 POTENTIAL.  
 FT TRANSMEM 207 227 POTENTIAL.  
 FT TRANSMEM 246 266 POTENTIAL.  
 FT TRANSMEM 283 303 POTENTIAL.  
 FT TRANSMEM 336 356 POTENTIAL.  
 FT TRANSMEM 361 381 POTENTIAL.  
 FT TRANSMEM 409 429 POTENTIAL.  
 FT TRANSMEM 435 455 POTENTIAL.  
 SQ SEQUENCE 467 AA: 51634 MW: 1152203134662 CRC64;

Query Match 79.5%; Score 35; DB 1; Length 467;  
 Best Local Similarity 80.0%; Pred. No. 65;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 1 WLYWI 5  
 98 WLYWL 102

Search completed: January 3, 2003, 15:29:20  
 Job time: 6.18182 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:24:35 ; Search time 20.3182 seconds  
(without alignments)  
60.846 Million cell updates/sec

Title: US-09-543-188a-4  
Perfect score: 44  
Sequence: 1 WLXWIP 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTRMBL\_21.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_prodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	95.5	3944	5	Q18667
2	40	90.9	309	11	Q9D029
3	40	90.9	533	4	Q14562
4	39	88.6	1564	5	Q9XM49
5	38	86.4	71	16	Q8XZU8
6	38	86.4	142	5	Q9V596
7	38	86.4	211	16	Q9I118
8	38	86.4	402	5	Q21056
9	38	86.4	402	5	Q9NEO0
10	38	86.4	564	3	Q9ETZ3
11	38	86.4	739	5	Q01260
12	38	86.4	739	16	Q98GK3
13	37	84.1	68	11	Q64622
14	37	84.1	322	13	Q9IBK4
15	37	84.1	340	11	Q99M14
16	37	84.1	385	5	Q9W3V0

17	37	84.1	406	11	Q922E7	Q922E7 mus musculus
18	37	84.1	427	10	Q93YN4	Q93YN4 arabidopsis
19	37	84.1	438	17	Q96XR2	Q96XR2 sulfobus
20	37	84.1	461	16	Q8YAS6	Q8YAS6 listeria mo
21	37	84.1	491	10	Q9FMR7	Q9FMR7 arabidopsis
22	37	84.1	525	3	Q74899	Q74899 schizosach
23	37	84.1	561	17	Q9HKV9	Q9HKV9 thermoplasm
24	37	84.1	1476	5	Q965D3	Q965D3 dictyosteli
25	37	84.1	1476	5	Q8ST66	Q8ST66 dictyosteli
26	37	84.1	1613	4	Q75581	Q75581 homo sapien
27	37	84.1	1613	11	Q88572	Q88572 mus musculus
28	37	84.1	2701	6	Q8WN96	Q8WN96 bos taurus
29	36	81.8	174	12	Q8OLP7	Q8OLP7 mamestra co
30	36	81.8	218	16	Q9CJRS	Q9CJRS pasteurella
31	36	81.8	252	4	Q9H4J9	Q9H4J9 homo sapien
32	36	81.8	298	16	Q8X953	Q8X953 escherichia
33	36	81.8	340	16	Q9RYT4	Q9RYT4 mycoplasma
34	36	81.8	688	9	Q9FZR2	Q9FZR2 deinococcus
35	36	81.8	719	16	Q8Y9S1	Q8Y9S1 listeria mo
36	36	81.8	1191	3	Q42765	Q42765 candida alb
37	35	79.5	55	16	Q9KTF6	Q9KTF6 vibrio chol
38	35	79.5	126	16	Q8Y42	Q8Y42 anabaena sp
39	35	79.5	136	16	Q92DY2	Q92DY2 listeria in
40	35	79.5	152	3	Q12302	Q12302 saccharomyc
41	35	79.5	161	8	Q35341	Q35341 podospora a
42	35	79.5	179	16	Q99Q08	Q99Q08 streptomyce
43	35	79.5	228	17	Q8ZV94	Q8ZV94 pyrobaculum
44	35	79.5	255	16	Q8Z962	Q8Z962 salmonella
45	35	79.5	260	16	Q8XTJ4	Q8XTJ4 ralstonia s

## ALIGNMENTS

RESULT 1  
Q18667 ID Q18667 PRELIMINARY; PRT; 3944 AA.  
AC Q18667;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE C47D12.1 protein.  
GN C47D12.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea.  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gajadasy S.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C. elegans: A platform for  
investigating biology."  
RT Science 282:2012-2018(1998).  
DR EMBL; 269902; CAA93765.1; -.  
DR InterPro: IPR002048; EF-hand.  
DR InterPro: IPR003151; FAT.  
DR InterPro: IPR003152; FAT.  
DR InterPro: IPR00403; P13\_P14\_kinase.  
DR Pfam; PF02259; FAT; 1.  
DR Pfam; PF02260; FATC; 1.  
DR Pfam; PF00454; P13\_P14\_kinase; 1.  
DR SMART; SM00146; P13Kc; 1.  
DR PROSITE; PS00018; EF\_HAND; UNKNOWN.1.  
SQ SEQUENCE 3944 AA; 452752 MW; AD1B1971F16D923 CRC64;  
Query Match 95.5%; Score 42; DB 5; Length 3944;  
Best Local Similarity 83.3%; Pred. No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLYWIP 6  
 DB 3252 WLYWIP 3257

## RESULT 2

Q9D029 PRELIMINARY: PRT: 309 AA.  
 AC Q9D029:  
 DT 01-JUN-2001 (TRENBLrel. 17, Created)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
 DE DNA segment, Chr 7, Wayne state University 128, expressed (Unknown)  
 DE (Protein for MGC:19443).  
 GN D7WSU128.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxId=10090;  
 [1]

SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRIO;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima M., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Rung B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:685-690(2001).  
 [2]

RA SEQUENCE FROM N.A.  
 RP Strausberg R.;  
 RI Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RA EMBL: AK011862; BAB2786.1; -;  
 RA EMBL: BC011313; AAH11313.1; -;  
 DR MGD: MGI:107301; D7WSU128.  
 DR InterPro: IPR000626; Ubiquitin.  
 DR Pfam: PF00240; ubiquitin; 1.  
 DR SMART: SM00213; UBO; 1.  
 DR PROSITE: PS00299; UBIQUITIN\_1; UNKNOWN\_1.  
 DR PROSITE: PS50053; UBIQUITIN\_2; 1.  
 SQ SEQUENCE 309 AA; 33449 MW; 9063FACBAE5A9C8E CRC64;

Query Match 90.9%; Score 40; DB 11; Length 309;  
 Best Local Similarity 66.7%; Pred. No. 53;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLYWIP 6  
 DB 286 WLYWIP 291

RESULT 3  
 O14562 PRELIMINARY: PRT: 533 AA.  
 AC O14562:  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
 DE Gene product with similarity to ubiquitin binding enzyme.  
 GN A-73566.2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxId=9606;  
 [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=99425270; PubMed=10493829;  
 RA Loftus B.J., Kim U.T., Sneddon V.P., Kalush F., Brandon R.,  
 RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,  
 RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.L., Mitchell S.,  
 RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;  
 RT "Genome duplications and other features in 12 Mb of DNA sequence from  
 human chromosome 16p and 16q";  
 RL Genomics 60:285-308(1999).  
 DR EMBL: AC002400; AAC05812.1; -;  
 DR InterPro: IPR000626; Ubiquitin.  
 DR Pfam: PF00240; ubiquitin; 1.  
 DR SMART: SM00213; UBO; 1.  
 DR PROSITE: PS00299; UBIQUITIN\_1; UNKNOWN\_1.  
 DR PROSITE: PS50053; UBIQUITIN\_2; 1.  
 SQ SEQUENCE 533 AA; 56895 MW; 500B38770EC03595 CRC64;

Query Match 90.9%; Score 40; DB 4; Length 533;  
 Best Local Similarity 66.7%; Pred. No. 86;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLYWIP 6  
 DB 510 WLYWIP 515

RESULT 4  
 Q9XW49 PRELIMINARY: PRT: 1564 AA.  
 AC Q9XW49:  
 DT 01-NOV-1999 (TRENBLrel. 12, Created)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Y53C10A.9 protein.  
 GN Y53C10A.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 NCBI\_TaxId=6239;  
 [1]

RP SEQUENCE FROM N.A.  
 RP White S.;  
 RI Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 RA EMBL: AL033536; CA22142.1; -;  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR001623; ABC\_tran.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR Prodom: PD00006; ABC\_transporter; 2.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; UNKNOWN\_2.  
 DR PROSITE: PS00636; DNAU\_1; UNKNOWN\_1.  
 SQ SEQUENCE 1564 AA; 176701 MW; DC66A37ED8DC50F CRC64;

Query Match 88.6%; Score 39; DB 5; Length 1564;  
 Best Local Similarity 66.7%; Pred. No. 3.3e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLYWIP 6  
 DB 1 WLYWIP 6



Db 326 WLFWVP 331

RESULT 5

08XZ08 PRELIMINARY: PRT: 71 AA.

DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, last sequence update)

DE Putative transmembrane protein.

GN RSC1297 OR R502819.

OS Ralstonia solanacearum (Pseudomonas solanacearum).

OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;

OC Ralstonia.

OX NCBI\_TaxID=305;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=GM11000;

MEDLINE=21681879; PubMed=11823852;

Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,

Arlat M., Billault A., Brottier P., Camu J.C., Catolico L.,

Chandler M., Choisine N., Claudel-Renard C., Cunac S., Demange N.,

Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,

Siglier P., Thebaud P., Whalen M., Wincker P., Levy M.,

Weissenbach J., Boucher C.A.;

RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";

RT Nature 415:497-502(2002).

RL EMBL: AL646063; CAD14999.1; -.

KW Complete proteome.

SO SEQUENCE 71 AA; 8263 MW; F10632B9CF260C49 CRC64;

Query Match 86.4%; Score 38; DB 16; Length 71;

Best Local Similarity 66.7%; Pred. No. 28;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLYWIP 6

Db 44 WLFWLP 49

RESULT 6

Q9V596 PRELIMINARY: PRT: 142 AA.

AC Q9V596;

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)

GN CG12929 protein.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Gelinkler S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Chame M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,

Abra J.J.F., Adzhani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

BA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borovoy D., Botchan D., Butler H., Cadieu E., Center A., Chandra I.,

RA Burlis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegami C.,

RA Jafari M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Moharry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon S., Nusskern D.R., Paclet J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL: AE003833; AAF58921.1; -.

DR Flybase: FBgn0033429; CG12929.

SO SEQUENCE 142 AA; 15642 MW; 20DEA503BDC12DB CRC64;

Query Match 86.4%; Score 38; DB 5; Length 142;

Best Local Similarity 83.3%; Pred. No. 53;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLYWIP 6

Db 53 WRYWIP 58

RESULT 7

091118 PRELIMINARY: PRT: 211 AA.

AC 091118;

DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)

DT 01-OCT-2001 (TREMBlrel. 18, last annotation update)

DE Hypothetical protein PA2287.

GN PA2287.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI\_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 15692 / PA01;

MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,

RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagron M.,

RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yan X.,

Brody L.L., Coulter S.N., Folger K.R., Kas A., Lartig K., Lim R.M.,

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

opportunistic pathogen.";

RL Nature 406:959-964(2000).

DR EMBL: AE004654; AAG05673.1; -.

DR Hypothetical protein: Complete proteome.

SO SEQUENCE 211 AA; 23043 MW; 201B4710FE84AB59 CRC64;

Query Match 86.4%; Score 38; DB 16; Length 211;

Best Local Similarity 66.7%; Pred. No. 76;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLYWIP 6

Db 11 WLYWIP 6

Db 59 WLFMLP 64

## RESULT 8

Q21056 PRELIMINARY; PRT; 402 AA.  
AC Q21056;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
GN Fatty acid desaturase.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA "The fat-1 gene of Caenorhabditis elegans encodes an omega-3 fatty acid desaturase."  
KL Submitted (Mar-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL: L41807; AAA67369.1; -.  
DR InterPro: IPR001225; FA\_desaturase.  
DR Pfam: PF00487; FA\_desaturase; 1.  
DR ProDom: PD001081; FA\_desaturase; 2.  
SQ SEQUENCE 402 AA; 46545 MW; A26ADAC6F05B0F0 CRC64;

Query Match 86.4%; Score 38; DB 5; Length 402;  
Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLYWP 6  
Db 261 WYXWP 266

## RESULT 9

Q9NE00 PRELIMINARY; PRT; 402 AA.  
AC Q9NE00;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
GN Y67H2A, a protein (Y67H2A.8 protein).  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sulston J.E.;  
KL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL132951; CAC44309.1; -.  
DR InterPro: IPR001225; FA\_desaturase.  
DR Pfam: PF00487; FA\_desaturase; 1.  
DR ProDom: PD001081; FA\_desaturase; 2.  
SQ SEQUENCE 402 AA; 46559 MW; BE1B1CD794EA98C5 CRC64;

Query Match 86.4%; Score 38; DB 5; Length 402;

Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLYWP 6  
Db 261 WYXWP 266

## RESULT 10

O96T23 PRELIMINARY; PRT; 564 AA.  
AC O96T23;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
GN Hypothetical 62.9 kDa protein.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Allyn V., Hohnsels J., Brandt P., Fartmann B., Holland R.,  
RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
KL Submitted (Dec-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL451014; CAD11346.1; -.  
DR InterPro: IPR001138; Fungl\_TTN.  
DR Pfam: PF001172; Zn\_Clus; 1.  
DR PROSITE: PS00463; ZN2\_CY6\_FUNGAL\_1; UNKNOWN\_1.  
DR PROSITE: PS50048; ZN2\_CY6\_FUNGAL\_2; 1.  
KW DNA-binding; Hypothetical protein; Metal-binding; Nuclear protein;  
KW Transcription regulation; Zinc.  
SQ SEQUENCE 564 AA; 62860 MW; A7054767236C3CB CRC64;

Query Match 86.4%; Score 38; DB 3; Length 564;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLYWP 6  
Db 329 WLFMLP 334

## RESULT 11

O01260 PRELIMINARY; PRT; 739 AA.  
AC O01260;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
GN T20D3.11 protein.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lloyd C.R.;  
KL Submitted (Dec-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL132951; CAC44309.1; -.  
DR InterPro: IPR001225; FA\_desaturase.  
DR Pfam: PF00487; FA\_desaturase; 1.  
DR ProDom: PD001081; FA\_desaturase; 2.  
SQ SEQUENCE 739 AA; 80559 MW; BE1B1CD794EA98C5 CRC64;

Query Match 86.4%; Score 38; DB 5; Length 739;

RL Science 282:2012-2018(1998).  
DR EMBL: 268220; CAA92491.1; -  
SQ SEQUENCE 739 AA; 85116 MW; 47363290D1510329 CRC64;

Query Match 86.4%; Score 38; DB 5; Length 739;  
Best Local Similarity 83.3%; Pred. No. 2.4e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLYWIP 6  
| | | | |  
Db 475 WSYWIP 480

RESULT 12  
098GK3

ID 098GK3 PRELIMINARY; PRT: 739 AA.

AC 098GK3:  
DT 01-OCT-2001 (TREMBLrel. 18, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
Hypothetical protein ml13287.

OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.

OX NCBI\_TaxID=381;  
[1]

RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
MEDLINE=21082930; PubMed=11214968;

RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
Watanabe A., Idehara K., Ishikawa K., Kawashima K., Kimura T.,  
Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,  
Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
Mesorhizobium loti."  
RL DNA Res. 7:331-338(2000).  
DR EMBL: AP003001; BAB50213.1; -  
DR InterPro: IPR001880; MSion\_channel.  
DR Pfam: PF00924; MS\_channel; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 739 AA; 81423 MW; E57C8EB5A704087 CRC64;

Query Match 86.4%; Score 38; DB 16; Length 739;  
Best Local Similarity 86.7%; Pred. No. 2.4e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

● 1 WLYWIP 6  
| | | | |  
Db 259 WLYWIP 264

RESULT 13  
064622

ID 064622 PRELIMINARY; PRT: 68 AA.

AC 064622:  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Protein-tyrosine-phosphatase (EC 3.1.3.48) (Fragment).

GN PTP-S.

OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;  
[1]

RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
MEDLINE=96125184; PubMed=8534367;

RA Reddy R.S., Swatup G.;

RT "Alternative splicing generates four different forms of a non-membrane protein tyrosine phosphatase mRNA.";

RL DNA Cell Biol. 14:1007-1015(1995).

DR EMBL: X92747; CAA63406.1; -

DR Hydrolase.

FT NON\_TER

SQ SEQUENCE 68 AA; 8194 MW; DA5293915CABCEFCRC64;

Query Match 84.1%; Score 37; DB 11; Length 68;  
Best Local Similarity 83.3%; Pred. No. 39;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLYWIP 6  
| | | | |  
Db 35 WLYWIP 40

RESULT 14  
091BA4

ID 091BA4 PRELIMINARY; PRT: 322 AA.

AC 091BA4:  
DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE RYTPN3 protein (Fragment).

GN RYTPN3.

OS Potamotrygon motoro (South American freshwater stingray).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;

OC Elasmobranchii; Squalae; Hypnosquales; Pristigasterae; Batoidae;

OC Myliobatiformes; Myliobatoidae; Potamotrygonidae; Potamotrygon.

OX NCBI\_TaxID=86373;  
[1]

RP SEQUENCE FROM N.A.  
RC MEDLINE=20219325; PubMed=10754074;

RA Ono-Koyanagi K., Suga H., Kato K., Miyata T.;

RT "Protein tyrosine phosphatases from amphioxus, hagfish, and ray:  
divergence of tissue-specific isoform genes in the early evolution of  
vertebrates.";

RL J. Mol. Evol. 50:302-311(2000).

DR EMBL: AB033582; BAA95189.1; -

DR HSSP: P18031; 1B2J

DR InterPro: IPR000387; TYR\_phosphatase.

DR InterPro: IPR000242; TYR\_PP.

DR Pfam: PF00102; Y\_phosphatase; 1.

DR PRINTS: PR00700; PRTYHPHPTASE.

DR SMART: SM00194; PTPC; 1.

DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; 1.

DR PROSITE: PS50056; TYR\_PHOSPHATASE\_2; 1.

DR PROSITE: PS50055; TYR\_PHOSPHATASE\_PTP; 1.

KW Hydrolase.

FT NON\_TER

SQ SEQUENCE 322 AA; 36954 MW; B0045D23B66F1197 CRC64;

Query Match 84.1%; Score 37; DB 13; Length 322;  
Best Local Similarity 83.3%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLYWIP 6  
| | | | |  
Db 293 WLYWIP 298

RESULT 15  
099M14

ID 099M14 PRELIMINARY; PRT: 340 AA.

AC 099M14:  
DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Cathepsin S.

GN CTSS.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC002125; AH02125.1; -.  
 DR HSSP: P25774; IBBF.  
 DR MEROPS: C01.034; -.  
 DR MGD: MGI:107341; Ctss.  
 DR InterPro: IPR000668; Peptidase\_C1.  
 DR InterPro: IPR000169; SHProt\_acsite.  
 DR Pfam: PF00112; Peptidase\_C1; 1.  
 DR PRINTS: PR00705; PAPA1N.  
 DR ProDom: PD000158; Peptidase\_C1; 1.  
 DR PROSITE: PS00640; THIOL\_PROTEASE\_ASN; 1.  
 DR PROSITE: PS00139; THIOL\_PROTEASE\_CYS; 1.  
 DR PROSITE: PS00639; THIOL\_PROTEASE\_HIS; 1.  
 KW Hydrolase; Thiol protease.  
 SQ SEQUENCE 340 AA; 38456 MW; 076502611F319DB7 CRC64;

Query Match 84.1%; Score 37; DB 11; Length 340;  
 Fast Local Similarity 66.7%; Pred No. 1.7e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLYWIP 6  
 ||:|:|  
 DB 11 WLFWMP 16

Search completed: January 3, 2003, 15:31:58  
 Job time : 22.3182 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:16 ; Search time 9 seconds  
(without alignments)  
19.615 Million cell updates/sec

Title: US-09-543-188A-4

Perfect score: 44

Sequence: 1 WLWIP 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

al number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PC10S.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	86.4	402	US-09-025-578-2	Sequence 2, Appl1
2	38	86.4	402	US-09-330-235-2	Sequence 2, Appl1
3	34	77.3	10	US-08-556-597-159	Sequence 159, Appl
4	34	77.3	17	US-08-190-788A-233	Sequence 233, App
5	34	77.3	17	US-08-383-474B-236	Sequence 236, App
6	34	77.3	17	US-08-465-391A-233	Sequence 233, App
7	34	77.3	17	US-08-464-538B-233	Sequence 233, App
8	34	77.3	17	US-08-463-076E-287	Sequence 287, App
9	34	77.3	257	US-09-372-422A-28	Sequence 28, Appl
10	34	77.3	272	US-09-372-422A-26	Sequence 26, Appl
11	34	77.3	511	US-08-991-677-4	Sequence 4, Appl1
12	34	77.3	823	US-07-745-206A-15	Sequence 15, Appl
13	34	77.3	823	US-08-311-363-15	Sequence 15, Appl
14	34	77.3	1754	US-07-745-206A-13	Sequence 13, Appl
15	34	77.3	1754	US-08-311-363-13	Sequence 13, Appl
16	34	77.3	2237	US-08-455-543A-48	Sequence 48, Appl
17	34	77.3	2237	US-08-223-305C-48	Sequence 48, Appl
18	34	77.3	2237	US-09-268-163-8	Sequence 8, Appl1
19	34	77.3	2265	US-08-149-097D-36	Sequence 36, Appl
20	34	77.3	2336	US-09-268-163-10	Sequence 10, Appl
21	34	77.3	2337	US-08-713-118-2	Sequence 2, Appl1
22	34	77.3	2337	US-09-452-007-2	Sequence 2, Appl1
23	34	77.3	2339	US-08-455-543A-47	Sequence 47, Appl
24	34	77.3	2339	US-08-223-305C-47	Sequence 47, Appl
25	34	77.3	2339	US-09-268-163-6	Sequence 6, Appl1
26	34	77.3	2343	US-09-268-163-4	Sequence 4, Appl1
27	34	77.3	2509	US-08-149-097D-35	Sequence 35, Appl1

28	33	75.0	118	US-09-134-001C-4174	Sequence 4174, Ap
29	33	75.0	134	US-09-273-163-8	Sequence 8, Appl1
30	33	75.0	247	US-09-372-422A-48	Sequence 48, Appl
31	33	75.0	249	US-09-372-422A-22	Sequence 22, Appl
32	33	75.0	249	US-09-372-422A-30	Sequence 30, Appl
33	33	75.0	249	US-09-372-422A-6	Sequence 6, Appl1
34	33	75.0	250	US-09-372-422A-24	Sequence 24, Appl
35	33	75.0	254	US-09-372-422A-34	Sequence 34, Appl
36	33	75.0	576	US-08-864-785-1	Sequence 1, Appl1
37	33	75.0	608	US-09-134-001C-3395	Sequence 3395, Ap
38	33	75.0	674	US-08-803-973-7	Sequence 7, Appl1
39	33	75.0	674	US-08-803-973-7	Sequence 7, Appl1
40	33	75.0	707	US-08-803-973-12	Sequence 12, Appl
41	33	75.0	707	US-08-803-973-12	Sequence 12, Appl
42	33	75.0	746	US-08-476-519-11	Sequence 11, Appl
43	33	75.0	746	PCT-US95-09323-11	Sequence 11, Appl
44	33	75.0	777	US-08-476-519-2	Sequence 2, Appl1
45	33	75.0	777	PCT-US95-09323-2	Sequence 2, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-025-578-2  
Sequence 2, Application US/09025578  
Patent No. 6194167  
GENERAL INFORMATION:  
APPLICANT: John A. Browse and James P. Spychalla  
TITLE OF INVENTION: OMEGA-3 FATTY ACID DESATURASE  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Alan E. Dow, Ph.D.  
STREET: One World Trade Center  
STREET: 121 S.W. Salmon Street  
STREET: Suite 1600  
CITY: Portland  
STATE: Oregon  
COUNTRY: United States of America  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Disk, 3-1/2 inch  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/025,578  
FILING DATE: Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/038,409  
FILING DATE: February 18, 1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Alan E. Dow, Ph.D.  
REGISTRATION NUMBER: 35,123  
REFERENCE/DOCKET NUMBER: 4630-49462/AED  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 226-7391  
TELEFAX: (503) 228-9446  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acid residues  
TYPE: amino acid  
STRANDEDNESS: single stranded  
TOPOLOGY: linear  
US-09-025-578-2  
Query Match 86.4%; Score 38; DB 4; Length 402;  
Best Local Similarity 66.7%; Pred. No. 65;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 WLWIP 6

Db 261 WYMWVP 266

RESULT 2  
US-09-330-235-2  
; Sequence 2, Application US/09330235  
; Patent No. 6459018  
; GENERAL INFORMATION:  
; APPLICANT: Knutzen, Debbie  
; TITLE OF INVENTION: POLYUNSATURATED FATTY ACIDS IN PLANTS  
; FILE REFERENCE: MOCO.156.0005  
; CURRENT APPLICATION NUMBER: US/09/330,235  
; PRIOR FILING DATE: 1999-06-10  
; PRIOR APPLICATION NUMBER: 60/089,043  
; PRIOR FILING DATE: 1998-06-12  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 402  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-09-330-235-2

Query Match 86.4%; Score 38; DB 4; Length 402;  
Best Local Similarity 66.7%; Pred. No. 65;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLYWIP 6  
1 1 1 1  
Db 261 WYMWVP 266

RESULT 3  
US-08-556-597-159  
; Sequence 159, Application US/08556597  
; Patent No. 5877155  
; GENERAL INFORMATION:  
; APPLICANT: Miller, Jonathan L.  
; TITLE OF INVENTION: MIMOTOPES AND ANTI-MIMOTOPES OF  
; TITLE OF INVENTION: HUMAN PLATELET GLYCOPROTEIN ID/IX  
; NUMBER OF SEQUENCES: 173  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devens & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/556,597  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/406,330  
; FILING DATE: 17-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Timian, Susan J.  
; REGISTRATION NUMBER: 34,103  
; REFERENCE/DOCKET NUMBER: 20884/101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1636  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 159:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid.

; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-556-597-159

Query Match 77.3%; Score 34; DB 2; Length 10;  
Best Local Similarity 50.0%; Pred. No. 7.8;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLYWIP 6  
1 1 1 1  
Db 3 WYMWVP 8

RESULT 4  
US-08-190-788A-233  
; Sequence 233, Application US/08190788A  
; Patent No. 5608035  
; GENERAL INFORMATION:  
; APPLICANT: Yanofsky, Stephen D.  
; APPLICANT: Barrett, Ronald W.  
; APPLICANT: Baldwin, David N.  
; APPLICANT: Jacobs, Jeff W.  
; TITLE OF INVENTION: Peptides and Compounds That Bind to the  
; TITLE OF INVENTION: IL-1 Receptor  
; NUMBER OF SEQUENCES: 312  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Affymax Technologies N.V.  
; STREET: 4001 Miranda Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/190,788A  
; FILING DATE: 02-FEB-1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847,567  
; FILING DATE: 05-MAR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stevens, Lauren L.  
; REGISTRATION NUMBER: 36,691  
; REFERENCE/DOCKET NUMBER: 1019.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-496-2300  
; TELEFAX: 415-424-0832  
; INFORMATION FOR SEQ ID NO: 233:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-190-788A-233

Query Match 77.3%; Score 34; DB 1; Length 17;  
Best Local Similarity 66.7%; Pred. No. 13;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLYWIP 6  
1 1 1 1  
Db 4 WYMWVP 9

RESULT 5  
US-08-383-474B-236  
; Sequence 236, Application US/08383474B

```

; Patent No. 5767234
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Stephen D.
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Baldwin, David N.
; APPLICANT: Jacobs, Jeff W.
; TITLE OF INVENTION: Peptides and Compounds That Bind to
; TITLE OF INVENTION: the IL-1 Receptor
; NUMBER OF SEQUENCES: 314
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend & Townsend & Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/383,474B
; FILING DATE: 01-FEB-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/190,788
; FILING DATE: 02-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens, Lauren L.
; REGISTRATION NUMBER: 36,691
; REFERENCE/DOCKET NUMBER: 1019.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-496-2300
; TELEFAX: 415-424-0832
; INFORMATION FOR SEQ ID NO: 236:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-383-474B-236

Query Match          77.3%; Score 34; DB 1; Length 17;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 WLWIP 6
1 ||| |
4 WYWQP 9

RESULT 6
US-08-465-391A-233
; Sequence 233, Application US/08465391A
; Patent No. 5766331
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Yanofsky, Stephen D.
; APPLICANT: Baldwin, David
; APPLICANT: Jacobs, Jeff W.
; APPLICANT: Boyy, Philippe R.
; APPLICANT: Leahy, Ellen M.
; APPLICANT: Portorff, Richard S.
; TITLE OF INVENTION: Peptides and Compounds That Bind to the
; TITLE OF INVENTION: IL-1 Receptor
; NUMBER OF SEQUENCES: 405
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
```

```

; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,391A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,474
; FILING DATE: 01-FEB-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/190,788
; FILING DATE: 02-FEB-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5786331v1el, Vern
; REGISTRATION NUMBER: 32,483
; REFERENCE/DOCKET NUMBER: 16528A-001840/1019.2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 233:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-465-391A-233

Query Match          77.3%; Score 34; DB 1; Length 17;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLWIP 6
1 ||| |
Db 4 WYWQP 9

RESULT 7
US-08-464-538B-233
; Sequence 233, Application US/08464538B
; Patent No. 5861476
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Yanofsky, Stephen D.
; APPLICANT: Baldwin, David
; APPLICANT: Jacobs, Jeff W.
; APPLICANT: Boyy, Philippe R.
; APPLICANT: Leahy, Ellen M.
; APPLICANT: Portorff, Richard S.
; TITLE OF INVENTION: Peptides and Compounds That Bind to the
; TITLE OF INVENTION: IL-1 Receptor
; NUMBER OF SEQUENCES: 402
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/464,538B  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/373,474  
FILING DATE: 01-FEB-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/190,788  
FILING DATE: 02-FEB-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, William M.  
REGISTRATION NUMBER: 30,223  
REFERENCE/DOCKET NUMBER: 16528A-001810  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 233:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-464-538B-233

Query Match 77.3%; Score 34; DB 2; Length 17;  
Best local Similarity 66.7%; Pred. No. 13;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLXWIP 6  
I:| | |  
Db 4 WYVWQP 9

RESULT 8  
US-08-463-076E-287  
Sequence 287, Application US/08463076E  
Patent No. 5880096  
GENERAL INFORMATION:  
APPLICANT: Barlett, Ronald W.  
APPLICANT: Yanofsky, Stephen D.  
TITLE OF INVENTION: Peptides and Compounds That Bind to the  
TITLE OF INVENTION: IL-1 Receptor  
NUMBER OF SEQUENCES: 392  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,076E  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Snyder, Joseph R.  
REGISTRATION NUMBER: 39,381  
REFERENCE/DOCKET NUMBER: 16528A-001850US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 287:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid.

STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-463-076E-287

Query Match 77.3%; Score 34; DB 2; Length 17;  
Best local Similarity 66.7%; Pred. No. 13;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLXWIP 6  
I:| | |  
Db 4 WYVWQP 9

RESULT 9  
US-09-372-422A-28  
Sequence 28, Application US/09372422A  
Patent No. 6313375  
GENERAL INFORMATION:  
APPLICANT: Rudolf Jung  
APPLICANT: Francois Barlieu  
TITLE OF INVENTION: Maize Aquaporins and Uses Thereof  
FILE REFERENCE: 0919  
CURRENT APPLICATION NUMBER: US/09/372,422A  
CURRENT FILING DATE: 1999-08-11  
PRIOR APPLICATION NUMBER: US 60/098,692  
PRIOR FILING DATE: 1998-08-31  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 28  
LENGTH: 257  
TYPE: PRT  
ORGANISM: Zea mays  
US-09-372-422A-28

Query Match 77.3%; Score 34; DB 4; Length 257;  
Best local Similarity 80.0%; Pred. No. 1,7e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLXWI 5  
I:| | |  
Db 222 WYVWI 226

RESULT 10  
US-09-372-422A-26  
Sequence 26, Application US/09372422A  
Patent No. 6313375  
GENERAL INFORMATION:  
APPLICANT: Rudolf Jung  
APPLICANT: Francois Barlieu  
TITLE OF INVENTION: Maize Aquaporins and Uses Thereof  
FILE REFERENCE: 0919  
CURRENT APPLICATION NUMBER: US/09/372,422A  
CURRENT FILING DATE: 1999-08-11  
PRIOR APPLICATION NUMBER: US 60/098,692  
PRIOR FILING DATE: 1998-08-31  
NUMBER OF SEQ ID NOS: 49  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 26  
LENGTH: 272  
TYPE: PRT  
ORGANISM: Zea mays  
US-09-372-422A-26

Query Match 77.3%; Score 34; DB 4; Length 272;  
Best local Similarity 80.0%; Pred. No. 1,8e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLXWI 5  
I:| | |  
Db 220 WYVWI 224



## RESULT 11

US-08-991-677-4  
; Sequence 4, Application US/08991677A  
; Patent No. 6252135  
; GENERAL INFORMATION:  
; APPLICANT: Chiang, Vincent L  
; APPLICANT: Carraway, Daniel T  
; APPLICANT: Smeltzer, Richard H  
; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms  
; FILE REFERENCE: 50617  
; CURRENT APPLICATION NUMBER: US/08/991,677A  
; CURRENT FILING DATE: 1997-12-16  
; EARLIER APPLICATION NUMBER: US 60/033,381  
; EARLIER FILING DATE: 1996-12-16  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 511  
; TYPE: PRT  
; ORGANISM: Liquidambar styraciflua  
; US-08-991-677-4

Query Match 77.3%; Score 34; DB 4; Length 511;  
Best Local Similarity 66.7%; Pred. No. 3.4e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLXWIP 6  
|||:|  
DB 224 WLKWP 229

## RESULT 12

US-07-745-206A-15  
; Sequence 15, Application US/07745206A  
; Patent No. 5429921  
; GENERAL INFORMATION:  
; APPLICANT: Harpold, Michael  
; APPLICANT: Ellis, Steven  
; APPLICANT: Williams, Mark  
; APPLICANT: McCue, Ann  
; APPLICANT: Feldman, Daniel  
; TITLE OF INVENTION: Human Calcium Channel Compositions and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fitch, Even, Tabin & Flannery  
; STREET: 135 S. LaSalle  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: U.S.A.  
; ZIP: 60603

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/745,206A  
; FILING DATE: 19910815  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
; NAME: Feder, Scott B  
; REFERENCE/DOCKET NUMBER: 51504  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-372-7842

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:  
; LENGTH: 823 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-745-206A-15

Query Match 77.3%; Score 34; DB 1; Length 823;  
Best Local Similarity 83.3%; Pred. No. 5.3e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLXWIP 6  
|||:|  
DB 333 WLXWIP 338

## RESULT 13

US-08-311-363-15  
; Sequence 15, Application US/08311363  
; Patent No. 5876958  
; GENERAL INFORMATION:  
; APPLICANT: Harpold, Michael  
; APPLICANT: Ellis, Steven  
; APPLICANT: Williams, Mark  
; APPLICANT: Feldman, Daniel  
; APPLICANT: McCue, Ann  
; APPLICANT: Brenner, Robert  
; TITLE OF INVENTION: Human Calcium Channel Compositions and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brown, Martin, Haller & McClain  
; STREET: 1660 Union Street  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92101-2926

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/311,363  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/745,206  
; FILING DATE: 15-AUG-1991

ATTORNEY/AGENT INFORMATION:  
; NAME: Seidman, Stephanie L.  
; REGISTRATION NUMBER: 33,779  
; REFERENCE/DOCKET NUMBER: 6362-51506  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619)238-0999  
; TELEFAX: (619)238-0062

INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 823 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-311-363-15

Query Match 77.3%; Score 34; DB 2; Length 823;  
Best Local Similarity 83.3%; Pred. No. 5.3e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLXWIP 6  
|||:|  
DB 333 WLXWIP 338

## RESULT 14

US-07-745-206A-13  
; Sequence 13, Application US/07745206A  
; Patent No. 5429921  
; GENERAL INFORMATION:  
; APPLICANT: Harpold, Michael  
; APPLICANT: Ellis, Steven

APPLICANT: Williams, Mark  
APPLICANT: McCue, Ann  
APPLICANT: Feldman, Daniel  
TITLE OF INVENTION: Human Calcium Channel Compositions and  
TITLE OF INVENTION: Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Filch, Even, Tablin & Flannery  
STREET: 135 S. LaSalle  
CITY: Chicago  
STATE: Illinois  
COUNTRY: U.S.A.  
ZIP: 60603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/745,206A  
FILING DATE: 19910815  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Feder, Scott B  
REFERENCE/DOCKET NUMBER: 51504  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-372-7842  
INFORMATION FOR SEQ. ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1754 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-745-206A-13

Query Match 77.3%; Score 34; DB 1; Length 1754;  
Best Local Similarity 83.3%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLXWIP 6  
DB 333 WLXWIP 338

RESULT 15  
US-08-311-363-13  
Sequence 13, Application US/08311363  
Patent No. 5876958  
GENERAL INFORMATION:

APPLICANT: Harpold, Michael  
APPLICANT: Ellis, Steven  
APPLICANT: Williams, Mark  
APPLICANT: Feldman, Daniel  
APPLICANT: McCue, Ann  
APPLICANT: Brenner, Robert  
TITLE OF INVENTION: Human Calcium Channel Compositions and  
TITLE OF INVENTION: Methods  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brown, Martin, Haller & McClain  
STREET: 1660 Union Street  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-2926  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/311.363  
FILING DATE:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/745,206  
FILING DATE: 15-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Seidman, Stephanie L.  
REGISTRATION NUMBER: 33,779  
REFERENCE/DOCKET NUMBER: 6362-51506  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619)238-0999  
TELEFAX: (619)238-0062  
INFORMATION FOR SEQ. ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1754 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-311-363-13

Query Match 77.3%; Score 34; DB 2; Length 1754;  
Best Local Similarity 83.3%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLXWIP 6  
DB 333 WLXWIP 338

Search completed: January 3, 2003, 15:34:28  
Job time : 10 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:32:11 ; Search time 4.77273 Seconds  
(without alignments)  
23.825 Million cell updates/sec

Title: US-09-543-188A-4  
Perfect score: 44  
Sequence: 1 WLWIP 6

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 117078 seqs, 18951520 residues  
al number of hits satisfying chosen parameters: 117078

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PC1\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PC1US\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	86.4	402	9	US-10-104-339-2
2	38	86.4	402	10	US-09-747-755-2
3	36	81.8	456	10	US-09-801-368-204
4	35	79.5	83	10	US-09-864-761-47931
5	35	79.5	360	10	US-09-815-242-11186
6	35	79.5	475	10	US-09-815-242-10052
7	34	77.3	446	9	US-10-042-141-58
8	34	77.3	446	10	US-09-726-643-58
9	34	77.3	510	9	US-10-042-141-139
10	34	77.3	510	10	US-09-726-643-139
11	34	77.3	511	10	US-09-796-256A-4
12	34	77.3	627	10	US-09-862-027-46
13	34	77.3	1034	9	US-09-738-626-4785
14	34	77.3	1189	9	US-09-738-626-4140
15	34	77.3	2237	12	US-10-033-026-8
16	34	77.3	2336	12	US-10-033-026-10
17	34	77.3	2339	12	US-10-033-026-6
18	34	77.3	2343	12	US-10-033-026-4
19	33	75.0	31	9	US-10-016-157A-201

20	33	75.0	56	10	US-09-864-761-35892	Sequence 35892, A
21	33	75.0	88	10	US-09-764-853-591	Sequence 591, App
22	33	75.0	126	10	US-09-864-761-48848	Sequence 48848, A
23	33	75.0	221	10	US-09-925-300-1659	Sequence 1659, App
24	33	75.0	255	10	US-09-864-711-15	Sequence 15, App1
25	33	75.0	262	9	US-09-981-353-63	Sequence 63, App1
26	33	75.0	288	10	US-09-925-299-840	Sequence 840, App
27	33	75.0	365	10	US-09-815-242-13185	Sequence 13185, A
28	33	75.0	398	10	US-09-729-674-146	Sequence 146, App
29	33	75.0	463	9	US-09-738-626-6546	Sequence 6946, App
30	33	75.0	466	10	US-09-815-242-10265	Sequence 10265, A
31	33	75.0	466	10	US-09-815-242-14048	Sequence 14048, A
32	33	75.0	475	10	US-09-815-242-5055	Sequence 5055, App
33	33	75.0	475	10	US-09-815-242-11503	Sequence 11503, A
34	33	75.0	519	9	US-09-895-913A-118	Sequence 118, App
35	33	75.0	519	10	US-09-815-242-11388	Sequence 11388, A
36	33	75.0	559	10	US-09-820-721A-1	Sequence 1, App1
37	33	75.0	563	10	US-09-915-181A-6	Sequence 6, App1
38	33	75.0	777	10	US-09-815-242-4894	Sequence 4894, App
39	33	75.0	914	10	US-09-815-242-10897	Sequence 10897, A
40	32	72.7	1167	10	US-09-815-242-11522	Sequence 11522, A
41	32	72.7	70	10	US-09-864-761-46021	Sequence 46021, A
42	32	72.7	79	9	US-09-738-626-4898	Sequence 4898, App
43	32	72.7	110	10	US-09-867-550-298	Sequence 298, App
44	32	72.7	116	10	US-09-741-843-4	Sequence 4, App1
45	32	72.7	169	10	US-09-925-301-902	Sequence 902, App

## ALIGNMENTS

RESULT 1:  
US-10-104-339-2  
Sequence 2, Appli  
Patent No. US20020170090A1  
GENERAL INFORMATION:  
APPLICANT: John A. Browne and James P. Spycharla  
TITLE OF INVENTION: OMEGA-3 FATTY ACID DESATURASE  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Alan E. Dow, Ph.D.  
STREET: One World Trade Center  
Suite 1600  
CITY: Portland  
STATE: Oregon  
COUNTRY: United States of America  
ZIP: 97204  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Disk, 3-1/2 inch  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/104,339  
FILING DATE: 21-Mar-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/747,755  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Alan E. Dow, Ph.D.  
REGISTRATION NUMBER: 35,123  
REFERENCE/DOCKET NUMBER: 4630-4946/AED  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (503) 226-7391  
TELEFAX: (503) 228-9446  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acid residues  
TYPE: amino acid  
STRANDEDNESS: single stranded  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-104-339-2

Query Match 86.4% Score 38; DB 9; Length 402;  
Best Local Similarity 66.7% Pred. No. 39;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLYWIP 6  
| | | |  
DB 261 WYWWP 266

RESULT 2

US-09-747-755-2  
Sequence 2, Application US/09747755  
Patent No. US2002004293A1

GENERAL INFORMATION:

APPLICANT: John A. Browse and James P. Spychalla

TITLE OF INVENTION: OMEGA-3 FATTY ACID DESATURASE

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Alan E. Dow, Ph.D.

STREET: One World Trade Center

STREET: 121 S.W. Salmon Street

CITY: Portland

STATE: Oregon

COUNTRY: United States of America

ZIP: 97204

COMPUTER READABLE FORM:

MEDIUM TYPE: Disk, 3-1/2 inch

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/747,755

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/025,578

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Alan E. Dow, Ph.D.

REGISTRATION NUMBER: 35,123

REFERENCE/DOCKET NUMBER: 4630-49462/AED

TELECOMMUNICATION INFORMATION:

TELEPHONE: (503) 226-7391

TELEFAX: (503) 226-9446

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 402 amino acid residues

TYPE: amino acid

STRANDEDNESS: single stranded

TOPOLOGY: linear

US-09-747-755-2

Query Match 86.4% Score 38; DB 10; Length 402;  
Best Local Similarity 66.7% Pred. No. 39;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLYWIP 6  
| | | |  
DB 261 WYWWP 266

RESULT 3

US-09-801-368-204

Sequence 204, Application US/09801368

Patent No. US20020128250A1

GENERAL INFORMATION:

APPLICANT: Busby, Robert

APPLICANT: Call, Brian

APPLICANT: Hecht, Peter  
APPLICANT: Holtzman, Doug  
APPLICANT: Madden, Kevin  
APPLICANT: Maxon, Mary  
APPLICANT: Milne, Todd  
APPLICANT: No. US20020128250A1man, Thea  
APPLICANT: Royer, John  
APPLICANT: Salama, Sofie  
APPLICANT: Sherman, Amir  
APPLICANT: Silva, Jeff  
APPLICANT: Summers, Eric  
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
FILE REFERENCE: 109272.147  
CURRENT APPLICATION NUMBER: US/09/801,368  
CURRENT FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 09/487,558  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: US 60/160,587  
PRIOR FILING DATE: 1999-10-20  
NUMBER OF SEQ ID NOS: 440  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 204  
LENGTH: 456  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-801-368-204

Query Match 81.8% Score 36; DB 10; Length 456;  
Best Local Similarity 66.7% Pred. No. 86;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLYWIP 6  
| | | |  
DB 22 WYWWP 27

RESULT 4

US-09-864-761-47931

Sequence 47931, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

FILE REFERENCE: Aecm1ca-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 47931
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009402.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.2
; OTHER INFORMATION: EST_HUMAN HIT: BF090287.1, EVALU0 3.10e-01
; OTHER INFORMATION: SWISSPROT HIT: P13276, EVALU0 1.90e+00
; US-09-864-761-47931

Query Match
Best Local Similarity 79.5%; Score 35; DB 10; Length 83;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLXWI 5
Db 17 WLXWL 21

RESULT 5
US-09-815-242-11186
; Sequence 1186, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 11186
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
```

```

US-09-815-242-11186

Query Match
Best Local Similarity 79.5%; Score 35; DB 10; Length 360;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLXWI 5
Db 147 WLXWL 151

RESULT 6
US-09-815-242-10052
; Sequence 1052, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 10052
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-815-242-10052

Query Match
Best Local Similarity 79.5%; Score 35; DB 10; Length 475;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLXWI 5
Db 105 WLXWL 109

RESULT 7
US-10-042-141-58
; Sequence 58, Application US/10042141
; Publication No. US20020183503A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: P2040P1
; CURRENT APPLICATION NUMBER: US/10/042,141
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/726,643
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
```

;; PRIOR FILING DATE: 2000-06-02  
;; PRIOR APPLICATION NUMBER: 60/137,725  
;; PRIOR FILING DATE: 1999-06-07  
;; NUMBER OF SEQ ID NOS: 190  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO 58  
;; LENGTH: 446  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-042-141-58

Query Match                      77.3%;    Score 34;    DB 9;    Length 446;  
Best Local Similarity    80.0%;    Pred. No. 1.7e+02;  
Matches    4;    Conservative    1;    Mismatches    0;    Indels    0;    Gaps    0;

QY            1 WLYWI 5  
              |:|:|  
DB    141 WYWI 145

;; RESULT 8  
;; US-09-726-643-58

;; Sequence 58, Application US/09726643  
;; Patent No. US20020028449A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ruben et al.  
;; TITLE OF INVENTION: 26 Human secreted proteins  
;; FILE REFERENCE: P2040P1  
;; CURRENT APPLICATION NUMBER: US/09/726,643  
;; CURRENT FILING DATE: 2000-12-01  
;; PRIOR APPLICATION NUMBER: PCT/US00/15187  
;; PRIOR FILING DATE: 2000-06-02  
;; PRIOR APPLICATION NUMBER: 60/137,725  
;; PRIOR FILING DATE: 1999-06-07  
;; NUMBER OF SEQ ID NOS: 190  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO 58  
;; LENGTH: 446  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-726-643-58

Query Match                      77.3%;    Score 34;    DB 10;    Length 446;  
Best Local Similarity    80.0%;    Pred. No. 1.7e+02;  
Matches    4;    Conservative    1;    Mismatches    0;    Indels    0;    Gaps    0;

QY            1 WLYWI 5  
              |:|:|  
DB    141 WYWI 145

;; RESULT 9  
;; US-10-042-141-139

;; Sequence 139, Application US/10042141  
;; Publication No. US20020183503A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ruben et al.  
;; TITLE OF INVENTION: 26 Human secreted proteins  
;; FILE REFERENCE: P2040P1  
;; CURRENT APPLICATION NUMBER: US/10/042,141  
;; CURRENT FILING DATE: 2002-01-11  
;; PRIOR APPLICATION NUMBER: 09/726,643  
;; PRIOR FILING DATE: 2000-12-01  
;; PRIOR APPLICATION NUMBER: PCT/US00/15187  
;; PRIOR FILING DATE: 2000-06-02  
;; PRIOR APPLICATION NUMBER: 60/137,725  
;; PRIOR FILING DATE: 1999-06-07  
;; NUMBER OF SEQ ID NOS: 190  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO 139  
;; LENGTH: 510  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens

US-10-042-141-139

Query Match                      77.3%;    Score 34;    DB 9;    Length 510;  
Best Local Similarity    80.0%;    Pred. No. 1.9e+02;  
Matches    4;    Conservative    1;    Mismatches    0;    Indels    0;    Gaps    0;

QY            1 WLYWI 5  
              |:|:|  
DB    205 WYWI 209

;; RESULT 10  
;; US-09-726-643-139

;; Sequence 139, Application US/09726643  
;; Patent No. US20020028449A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ruben et al.  
;; TITLE OF INVENTION: 26 Human secreted proteins  
;; FILE REFERENCE: P2040P1  
;; CURRENT APPLICATION NUMBER: US/09/726,643  
;; CURRENT FILING DATE: 2000-12-01  
;; PRIOR APPLICATION NUMBER: PCT/US00/15187  
;; PRIOR FILING DATE: 2000-06-02  
;; PRIOR APPLICATION NUMBER: 60/137,725  
;; PRIOR FILING DATE: 1999-06-07  
;; NUMBER OF SEQ ID NOS: 190  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO 139  
;; LENGTH: 510  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-726-643-139

Query Match                      77.3%;    Score 34;    DB 10;    Length 510;  
Best Local Similarity    80.0%;    Pred. No. 1.9e+02;  
Matches    4;    Conservative    1;    Mismatches    0;    Indels    0;    Gaps    0;

QY            1 WLYWI 5  
              |:|:|  
DB    205 WYWI 209

;; RESULT 11  
;; US-09-796-256A-4

;; Sequence 4, Application US/09796256A  
;; Patent No. US20020078477A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Chiang, Vincent L  
;; APPLICANT: Carriaway, Daniel T  
;; APPLICANT: Smeltzer, Richard H  
;; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms  
;; FILE REFERENCE: 50617/c-3532.0  
;; CURRENT APPLICATION NUMBER: US/09/796,256A  
;; CURRENT FILING DATE: 2001-02-28  
;; PRIOR APPLICATION NUMBER: US60/033381  
;; PRIOR FILING DATE: 1996-12-16  
;; PRIOR APPLICATION NUMBER: 08/991677  
;; PRIOR FILING DATE: 1997-12-12  
;; NUMBER OF SEQ ID NOS: 11  
;; SOFTWARE: Patentln Ver. 2.1  
;; SEQ ID NO 4  
;; LENGTH: 511  
;; TYPE: PRT  
;; ORGANISM: Liquidambar styraciflua  
US-09-796-256A-4

Query Match                      77.3%;    Score 34;    DB 10;    Length 511;  
Best Local Similarity    66.7%;    Pred. No. 1.9e+02;  
Matches    4;    Conservative    1;    Mismatches    1;    Indels    0;    Gaps    0;

QY            1 WLYWI 6  
              |:|:|  
DB    224 WLYWI 229

RESULT 12  
US-09-862-027-46  
; Sequence 46, Application US/09862027  
; Patent No. US20020142428A1  
; GENERAL INFORMATION:  
; APPLICANT: Hodge, Martin R.  
; TITLE OF INVENTION: NO. US20020142428A1 Kinases and Uses Thereof  
; FILE REFERENCE: 35800/234862  
; CURRENT APPLICATION NUMBER: US/09/862,027  
; CURRENT FILING DATE: 2001-05-21  
; PRIOR APPLICATION NUMBER: US 09/345,473  
; PRIOR FILING DATE: 1999-06-30  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 46  
; LENGTH: 627  
; TYPE: PRT  
; ORGANISM: Mus musculus  
09-862-027-46

Query Match 77.3%; Score 34; DB 10; Length 627;  
Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WLYWIP 6  
Db 210 WLYWIP 215

RESULT 13  
US-09-738-626-4785  
; Sequence 4785, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENO, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: Patentin ver. 3.0  
; SEQ ID NO 4785  
; LENGTH: 1034  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4785

Query Match 77.3%; Score 34; DB 9; Length 1034;  
Best Local Similarity 80.0%; Pred. No. 3.5e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLYWIP 5  
Db 563 WLYWIP 567

RESULT 14  
US-09-738-626-4140  
; Sequence 4140, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENO, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: Patentin ver. 3.0  
; SEQ ID NO 4140  
; LENGTH: 1189  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4140

Query Match 77.3%; Score 34; DB 9; Length 1189;  
Best Local Similarity 66.7%; Pred. No. 3.9e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WLYWIP 6  
Db 247 WLYWIP 252

RESULT 15  
US-10-033-026-8  
; Sequence 8, Application US/10033026  
; Patent No. US20020147309A1  
; GENERAL INFORMATION:  
; APPLICANT: Schorge, Stephanie  
; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF  
; FILE REFERENCE: B105577000  
; CURRENT APPLICATION NUMBER: US/10/033,026  
; CURRENT FILING DATE: 2001-12-28  
; PRIOR APPLICATION NUMBER: US 09/268,163  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: US 60/077,901  
; PRIOR FILING DATE: 1998-03-13  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 2237  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-033-026-8

Query Match 77.3%; Score 34; DB 12; Length 2237;  
Best Local Similarity 83.3%; Pred. No. 6.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLYWIP 6  
Db 333 WLYWIP 338

Mon Jan 6 15:09:04 2003

us-09-543-188a-4.rapb

Page 6

Search completed: January 3, 2003, 15:52:05  
Job time : 5.77273 secs

---



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:16 ; Search time 9 seconds  
(without alignments)  
19.615 Million cell updates/sec

Title: US-09-543-188A-13

Perfect score: 24

Sequence: 1 LLLVIA 6

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

al number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	100.0	112	1	US-07-800-364B-15
2	24	100.0	173	4	US-09-134-001C-4275
3	24	100.0	509	3	US-08-688-988-40
4	24	100.0	630	4	US-09-134-001C-4615
5	23	95.8	387	4	US-09-134-001C-4298
6	23	95.8	503	4	US-08-740-223A-18
7	23	95.8	503	4	US-09-134-001C-3948
8	23	95.8	503	4	US-09-202-491-4
9	23	95.8	503	4	US-09-202-491-10
10	23	95.8	503	4	US-09-709-188-18
11	23	95.8	913	1	US-08-445-640-4
12	23	95.8	913	3	US-08-170-558-4
13	23	95.8	913	3	US-08-447-314-4
14	23	95.8	913	3	US-08-445-461-4
15	23	95.8	919	3	US-08-336-343A-2
16	23	95.8	1210	2	US-08-484-438-7
17	23	95.8	1210	2	US-08-475-035-4
18	22	91.7	30	4	US-08-448-489-7
19	22	91.7	88	1	US-08-091-569-21
20	22	91.7	88	1	US-08-203-676-21
21	22	91.7	88	2	US-08-822-238-21
22	22	91.7	126	4	US-09-288-143-71
23	22	91.7	128	6	5179198-1
24	22	91.7	128	6	5521296-1
25	22	91.7	177	6	US-08-465-794-17
26	22	91.7	177	3	US-09-049-813-17
27	22	91.7	295	4	US-08-937-067-2

28	22	91.7	323	6	5185254-4	Patent No. 5185254
29	22	91.7	344	2	US-08-475-634D-19	Sequence 19, Appl
30	22	91.7	353	4	US-09-134-001C-4325	Sequence 4325, Ap
31	22	91.7	359	1	US-07-662-005A-4	Sequence 4, Appl1
32	22	91.7	503	4	US-09-382-256-10	Sequence 10, Appl
33	22	91.7	503	4	US-09-395-115-10	Sequence 10, Appl
34	22	91.7	503	4	US-08-436-265-10	Sequence 10, Appl
35	22	91.7	503	4	US-09-679-187-10	Sequence 10, Appl
36	22	91.7	514	4	US-09-717-432-2	Sequence 2, Appl1
37	22	91.7	514	4	US-09-912-484-2	Sequence 2, Appl1
38	22	91.7	579	3	US-08-704-711A-1	Sequence 1, Appl1
39	22	91.7	579	4	US-09-521-220-1	Sequence 1, Appl1
40	22	91.7	582	3	US-08-704-711A-2	Sequence 2, Appl1
41	22	91.7	582	4	US-08-448-489-1	Sequence 1, Appl1
42	22	91.7	582	4	US-09-211-704A-9	Sequence 9, Appl1
43	22	91.7	582	4	US-09-521-220-2	Sequence 2, Appl1
44	22	91.7	582	4	US-09-391-104-28	Sequence 28, Appl
45	22	91.7	976	2	US-08-449-645A-18	Sequence 18, Appl

#### ALIGNMENTS

RESULT 1  
US-07-800-364B-15  
; Sequence 15, Application US/07800364B  
; Patent No. 5688678  
; GENERAL INFORMATION:  
; APPLICANT: Hewick, Rodney M.  
; APPLICANT: Wang, Jack H.  
; APPLICANT: Mooney, John M.  
; APPLICANT: Celeste, Anthony J.  
; TITLE OR INVENTION: Bone and Cartilage Inductive Proteins  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.  
; STREET: 87 Cambridgepark Drive  
; CITY: Cambridge  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/800,364B  
; FILING DATE: 26-NOV-1991  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kapinos, Ellen J.  
; REGISTRATION NUMBER: 32,245  
; REFERENCE/DOCKET NUMBER: GI 5182A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-876-1170  
; TELEFAX: 617-876-5851  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 112 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; US-07-800-364B-15

Query Match 100.0%; Score 24; DB 1; Length 112;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLLVIA 6

Db 26 LLLVIA 31  
|||||

RESULT 2  
US-09-134-001C-4275

; Sequence 4275, Application US/09134001C  
; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; FILE REFERENCE: GTC-007  
EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 4275

; LENGTH: 173

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4275

Query Match

Best Local Similarity 100.0%; Score 24; DB 4; Length 173;  
Pred. No. 1.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||

Db 28 LLLVIA 33

RESULT 3  
US-08-688-988-40

; Sequence 40, Application US/08688988B

; Patent No. 6096545

; GENERAL INFORMATION:

; APPLICANT: Lefebvre, Daniel D.

; APPLICANT: Malboobi, Mohammad A.

; TITLE OF INVENTION: PHOSPHATE STARVATION-INDUCIBLE PROTEINS

; FILE REFERENCE: PPL96-03

; CURRENT APPLICATION NUMBER: US/08/688,988B

; PRIOR FILING DATE: 1996-07-31

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 40

; LENGTH: 509

; TYPE: PRT

; ORGANISM: Hordeum vulgare

US-08-688-988-40

Query Match

Best Local Similarity 100.0%; Score 24; DB 3; Length 509;  
Pred. No. 5.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||

Db 7 LLLVIA 12

RESULT 4  
US-09-134-001C-4615

; Sequence 4615, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; FILE REFERENCE: GTC-007  
EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 4615

; LENGTH: 630

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4615

Query Match

Best Local Similarity 100.0%; Score 24; DB 4; Length 630;  
Pred. No. 7.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||

Db 133 LLLVIA 138

RESULT 5  
US-09-134-001C-4298

; Sequence 4298, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; FILE REFERENCE: GTC-007  
EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/134,001C

; PRIOR FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 4298

; LENGTH: 387

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4298

Query Match

Best Local Similarity 95.8%; Score 23; DB 4; Length 387;  
Pred. No. 7.2e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||

Db 3 LLLVIA 8

RESULT 6  
US-08-740-223A-18

; Sequence 18, Application US/08740223A

; Patent No. 6265564

; GENERAL INFORMATION:

; APPLICANT: Davis, et al.

; TITLE OF INVENTION: Expressed Ligand - Vascular

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Regeneron Pharmaceuticals, Inc.

STREET: 777 Old Saw Mill Road

CITY: Tarrytown

STATE: NY

COUNTRY: USA

ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/740,223A  
;; FILING DATE: 25-OCT-1996  
;; CLASSIFICATION: 536  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: USSN 60/022/999  
;; FILING DATE: 02-AUG-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Covert, Robert J  
;; REGISTRATION NUMBER: 36,108  
;; REFERENCE/DOCKET NUMBER: REG 333  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 914-345-7400  
;; TELEFAX: 914-345-7721  
;; INFORMATION FOR SEQ ID NO: 18:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 503 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: internal  
;; FEATURE:  
;; NAME/KEY: TIE ligand-4  
;; LOCATION: 1...503  
;; OTHER INFORMATION:  
US-08-740-223A-18

Query Match 95.8%; Score 23; DB 4; Length 503;  
Best Local Similarity 83.3%; Pred. No. 9.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
Db 12 LLLVIA 17

RESULT 7  
US-09-134-001C-3948  
;; Sequence 3948, Application US/09134001C  
;; Patent No. 6380370  
;; GENERAL INFORMATION:  
;; APPLICANT: Lynn Doucette-Stamm et al  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
;; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
;; FILE REFERENCE: GTC-007  
;; CURRENT APPLICATION NUMBER: US/09/134,001C  
;; PRIOR FILING DATE: 1998-08-13  
;; PRIOR APPLICATION NUMBER: US 60/064,964  
;; PRIOR FILING DATE: 1997-11-08  
;; PRIOR APPLICATION NUMBER: US 60/055,779  
;; PRIOR FILING DATE: 1997-08-14  
;; NUMBER OF SEQ ID NOS: 5674  
;; SEQ ID NO 3948  
;; LENGTH: 503  
;; TYPE: PRT  
;; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3948

Query Match 95.8%; Score 23; DB 4; Length 503;  
Best Local Similarity 83.3%; Pred. No. 9.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
Db 268 LLLVIA 273

RESULT 8  
US-09-202-491-4  
;; Sequence 4, Application US/09202491  
;; Patent No. 6432667  
;; GENERAL INFORMATION:  
;; FILE REFERENCE: REG 333-2

;; APPLICANT: Valenzuela et al.  
;; TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF  
;; FILE REFERENCE: REG330-K  
;; CURRENT APPLICATION NUMBER: US/09/202,491  
;; CURRENT FILING DATE: 1998-11-16  
;; EARLIER APPLICATION NUMBER: PCT/US97/10728  
;; EARLIER FILING DATE: 1997-06-19  
;; EARLIER APPLICATION NUMBER: 60/022,999  
;; EARLIER FILING DATE: 1996-08-02  
;; EARLIER APPLICATION NUMBER: 60/021,087  
;; EARLIER FILING DATE: 1996-07-02  
;; EARLIER APPLICATION NUMBER: 08/665,926  
;; EARLIER FILING DATE: 1996-06-19  
;; NUMBER OF SEQ ID NOS: 14  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO 4  
;; LENGTH: 503  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-202-491-4

Query Match 95.8%; Score 23; DB 4; Length 503;  
Best Local Similarity 83.3%; Pred. No. 9.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
Db 12 LLLVIA 17

RESULT 9  
US-09-202-491-10  
;; Sequence 10, Application US/09202491  
;; Patent No. 6432667  
;; GENERAL INFORMATION:  
;; APPLICANT: Valenzuela et al.  
;; TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF  
;; FILE REFERENCE: REG330-K  
;; CURRENT APPLICATION NUMBER: US/09/202,491  
;; CURRENT FILING DATE: 1998-11-16  
;; EARLIER APPLICATION NUMBER: PCT/US97/10728  
;; EARLIER FILING DATE: 1997-06-19  
;; EARLIER APPLICATION NUMBER: 60/022,999  
;; EARLIER FILING DATE: 1996-08-02  
;; EARLIER APPLICATION NUMBER: 60/021,087  
;; EARLIER FILING DATE: 1996-07-02  
;; EARLIER APPLICATION NUMBER: 08/665,926  
;; EARLIER FILING DATE: 1996-06-19  
;; NUMBER OF SEQ ID NOS: 14  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO 10  
;; LENGTH: 503  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-202-491-10

Query Match 95.8%; Score 23; DB 4; Length 503;  
Best Local Similarity 83.3%; Pred. No. 9.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
Db 12 LLLVIA 17

RESULT 10  
US-09-709-188-18  
;; Sequence 18, Application US/09709188  
;; Patent No. 6441137  
;; GENERAL INFORMATION:  
;; APPLICANT: Davis et al.  
;; TITLE OF INVENTION: Expressed Ligand - Vascular Intercellular Signaling Molecule  
;; FILE REFERENCE: REG 333-2

;; CURRENT APPLICATION NUMBER: US/09/709,188  
;; CURRENT FILING DATE: 2000-11-09  
;; PRIOR APPLICATION NUMBER: 08/740,223  
;; PRIOR FILING DATE: 1996-10-25  
;; NUMBER OF SEQ ID NOS: 30  
;; SOFTWARE: Patent version 3.1  
;; SEQ ID NO: 18  
;; LENGTH: 503  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-709-188-18

Query Match 95.8%; Score 23; DB 4; Length 503;  
Best Local Similarity 83.3%; Pred. No. 9.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
11111  
DB 12 LLLVIA 17

RESULT 11

US-08-445-640-4  
; Sequence 4, Application US/08445640  
; Patent No. 5709858  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Mark, Melanie R.  
; APPLICANT: Scadden, David T.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Baron, Will F.  
; TITLE OF INVENTION: Protein Tyrosine Kinases  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,640  
; FILING DATE: 22-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/170558  
; FILING DATE: 20-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/157563  
; FILING DATE: 23-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 854C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 913 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-445-640-4

Query Match 95.8%; Score 23; DB 1; Length 913;  
Best Local Similarity 83.3%; Pred. No. 1.8e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
11111  
DB 430 LLLVIA 435

RESULT 12

US-08-170-558-4  
; Sequence 4, Application US/08170558  
; Patent No. 6001621  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Mark, Melanie R.  
; APPLICANT: Scadden, David T.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Baron, Will F.  
; TITLE OF INVENTION: Protein Tyrosine Kinases  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/170,558  
; FILING DATE: 20-DEC-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/157563  
; FILING DATE: 23-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 854C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 913 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-170-558-4

Query Match 95.8%; Score 23; DB 3; Length 913;  
Best Local Similarity 83.3%; Pred. No. 1.8e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
11111  
DB 430 LLLVIA 435

RESULT 13

US-08-447-314-4  
; Sequence 4, Application US/08447314  
; Patent No. 6087144  
; GENERAL INFORMATION:  
; APPLICANT: Scadden, David T.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Baron, Will F.  
; TITLE OF INVENTION: Protein Tyrosine Kinases  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/447,314  
FILING DATE: 22-MAY-1995  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/170558  
FILING DATE: 20-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/157563  
FILING DATE: 23-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 854CID2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 913 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-447-314-4

Query Match  
Best Local Similarity 83.3%; Score 23; DB 3; Length 913;  
Pred. No. 1.8e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLLVIA 6  
|||:|  
DB 430 LLLIIA 435

RESULT 14  
US-08-445-461-4  
Sequence 4, Application US/08445461  
Patent No. 6096527  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J.  
APPLICANT: Mark, Melanie R.  
APPLICANT: Scadden, David T.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Baron, Will F.  
TITLE OF INVENTION: Protein Tyrosine Kinases  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,461  
FILING DATE: 22-MAY-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/170558  
FILING DATE: 20-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/157563  
FILING DATE: 23-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 854C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 913 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-445-461-4

Query Match  
Best Local Similarity 83.3%; Score 23; DB 3; Length 913;  
Pred. No. 1.8e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLLVIA 6  
|||:|  
DB 430 LLLIIA 435

RESULT 15  
US-08-336-343A-2  
Sequence 2, Application US/08336343A  
Patent No. 5677144  
GENERAL INFORMATION:  
APPLICANT: Ullrich, Axel  
APPLICANT: Alves, Frauke  
TITLE OF INVENTION: CCK-2, A No. 5677144e1 Receptor Tyrosine Kinase  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/336,343A  
FILING DATE: 08-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-065  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 919 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-336-343A-2

Query Match  
Best Local Similarity 83.3%; Score 23; DB 1; Length 919;  
Pred. No. 1.8e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LLEVIA 6  
| | : | |  
Db 430 LLEVIA 435

Search completed: January 3, 2003, 15:34:29  
Job time : 10 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: January 3, 2003, 15:32:11 : Search time 4.77273 Seconds  
(without alignments)  
23.825 Million cell updates/sec

Title: US-09-543-188A-13

Perfect score: 24  
Sequence: 1 LLVIA 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 117078 seqs, 18951520 residues

al number of hits satisfying chosen parameters: 117078

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	100.0	96	US-10-016-481-11	Sequence 11, App1
2	24	100.0	179	US-09-764-853-847	Sequence 847, App
3	24	100.0	179	US-09-764-853-847	Sequence 296, App
4	24	100.0	478	US-09-764-853-847	Sequence 225, App
5	24	100.0	487	US-09-764-853-847	Sequence 680, App
6	24	100.0	504	US-09-738-626-4035	Sequence 4035, App
7	24	100.0	713	US-09-978-295A-183	Sequence 183, App
8	24	100.0	713	US-09-978-295A-183	Sequence 183, App
9	24	100.0	713	US-09-978-295A-183	Sequence 183, App
10	24	100.0	713	US-09-978-295A-183	Sequence 183, App
11	24	100.0	713	US-10-052-586-416	Sequence 416, App
12	24	100.0	859	US-09-529-063-70	Sequence 179, App
13	23	95.8	27	US-09-984-245-179	Sequence 34868, A
14	23	95.8	47	US-10-001-870-116	Sequence 110, App
15	23	95.8	62	US-09-950-933A-40	Sequence 976, App
16	23	95.8	100	US-09-925-299-976	Sequence 315, App
17	23	95.8	292	US-09-886-055-315	Sequence 10468, A
18	23	95.8	318	US-09-815-242-10468	
19	23	95.8	318	US-09-815-242-10468	

20	23	95.8	326	US-09-888-615-105	Sequence 105, App
21	23	95.8	377	US-10-091-628-2	Sequence 2, App1
22	23	95.8	394	US-09-925-301-1388	Sequence 1388, App
23	23	95.8	513	US-09-738-626-6860	Sequence 6860, App
24	23	95.8	624	US-09-925-297-738	Sequence 738, App
25	23	95.8	913	US-09-223-490-4	Sequence 4, App1
26	23	95.8	1210	US-09-725-433-2	Sequence 2, App1
27	23	91.7	39	US-09-864-761-47820	Sequence 47820, A
28	22	91.7	69	US-09-864-761-37443	Sequence 37443, A
29	22	91.7	106	US-09-738-626-4602	Sequence 4602, App
30	22	91.7	124	US-09-893-737-262	Sequence 262, App
31	22	91.7	167	US-10-003-152-18	Sequence 18, App1
32	22	91.7	258	US-09-965-529-25	Sequence 25, App1
33	22	91.7	271	US-09-738-626-5557	Sequence 5557, App
34	22	91.7	295	US-09-934-483A-1	Sequence 1, App1
35	22	91.7	295	US-09-934-483A-5	Sequence 5, App1
36	22	91.7	301	US-09-836-602-6	Sequence 6, App1
37	22	91.7	316	US-09-836-602-4	Sequence 4, App1
38	22	91.7	318	US-09-886-055-101	Sequence 101, App
39	22	91.7	322	US-09-815-242-11790	Sequence 11790, A
40	22	91.7	324	US-09-912-020-245	Sequence 245, App
41	22	91.7	359	US-09-925-302-591	Sequence 591, App
42	22	91.7	362	US-09-712-363-184	Sequence 184, App
43	22	91.7	373	US-09-836-602-2	Sequence 2, App1
44	22	91.7	381	US-09-738-626-5049	Sequence 5049, App
45	22	91.7	417	US-09-815-242-10165	Sequence 10165, A

## ALIGNMENTS

RESULT 1  
US-10-016-481-11  
; Sequence 11, Application US/10016481  
; Patent No. US20020115610A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhou, Qun-Yong  
; APPLICANT: Ehler, Frederick  
; TITLE OF INVENTION: Prokinection Polypeptides, Related  
; FILE REFERENCE: P-UC 5016  
; CURRENT APPLICATION NUMBER: US/10/016,481  
; CURRENT FILING DATE: 2001-11-01  
; PRIOR APPLICATION NUMBER: 60/245,882  
; PRIOR FILING DATE: 2000-11-03  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Bombina variegata  
US-10-016-481-11

Query Match 100.0%; Score 24; DB 12; Length 96;  
Best Local Similarity 100.0%; Pred No. 76;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLVIA 6  
|||||  
Db 10 LLVIA 15

RESULT 2  
US-09-764-853-847  
; Sequence 847, Application US/09764853  
; Patent No. US20020090672A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P206  
; CURRENT APPLICATION NUMBER: US/09/764,853  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper

```

; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 847
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (164)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (178)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-847
```

```

Query Match
Best Local Similarity 100.0%; Score 24; DB 10; Length 179;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LLLVIA 6
DB 134 LLLVIA 139
```

```

RESULT 3
US-09-764-898-296
; Sequence 296, Application US/09764898
; Patent No. US2002090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1201
; CURRENT APPLICATION NUMBER: US/09/764,898
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 296
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (164)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (178)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-898-296
```

```

Query Match
Best Local Similarity 100.0%; Score 24; DB 10; Length 179;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LLLVIA 6
DB 134 LLLVIA 139
```

```

RESULT 4
US-09-764-898-225
; Sequence 225, Application US/09764898
; Patent No. US2002090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1201
```

```

; CURRENT APPLICATION NUMBER: US/09/764,898
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 225
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-898-225
```

```

Query Match
Best Local Similarity 100.0%; Score 24; DB 10; Length 478;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LLLVIA 6
DB 125 LLLVIA 130
```

```

RESULT 5
US-09-764-853-680
; Sequence 680, Application US/09764853
; Patent No. US2002090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 680
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-680
```

```

Query Match
Best Local Similarity 100.0%; Score 24; DB 10; Length 487;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LLLVIA 6
DB 134 LLLVIA 139
```

```

RESULT 6
US-09-738-626-4035
; Sequence 4035, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAMA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 248-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
```



PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 4035  
LENGTH: 504  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4035

Query Match 100.0%; Score 24; DB 9; Length 504;  
Best Local Similarity 100.0%; Pred. No. 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
DB 490 LLLVIA 495

RESULT 7  
US-09-978-295A-183  
Sequence 183, Application US/09978295A  
Patent No. US20020156006A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Bolstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gertslen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tunas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C11  
CURRENT APPLICATION NUMBER: US/09/978, 295A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641

PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21

PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/08582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/08579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/08580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573

PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

Query Match 100.0% Score 24; DB 9; Length 713;  
Best Local Similarity 100.0%; Pred. No. 5; Gaps 0;  
Matches 6; Conservative 0; Mismatches 0; Indels 0;

OY 1 LLLVIA 6  
DB 454 LLLVIA 459

RESULT 8  
US-09-978-697-183  
Sequence 183: Application US/09978697  
Patent No. US20020169284A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Bolstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gertsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James;  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630PIC27  
CURRENT APPLICATION NUMBER: US/09/978, 697  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13

1	PRIOR APPLICATION NUMBER: 60/078886
2	PRIOR FILING DATE: 1998-03-20
3	PRIOR APPLICATION NUMBER: 60/078936
4	PRIOR FILING DATE: 1998-03-20
5	PRIOR APPLICATION NUMBER: 60/078910
6	PRIOR FILING DATE: 1998-03-20
7	PRIOR APPLICATION NUMBER: 60/078939
8	PRIOR FILING DATE: 1998-03-20
9	PRIOR APPLICATION NUMBER: 60/079294
10	PRIOR FILING DATE: 1998-03-25
11	PRIOR APPLICATION NUMBER: 60/079656
12	PRIOR FILING DATE: 1998-03-26
13	PRIOR APPLICATION NUMBER: 60/079664
14	PRIOR FILING DATE: 1998-03-27
15	PRIOR APPLICATION NUMBER: 60/079689
16	PRIOR FILING DATE: 1998-03-27
17	PRIOR APPLICATION NUMBER: 60/079693
18	PRIOR FILING DATE: 1998-03-27
19	PRIOR APPLICATION NUMBER: 60/079728
20	PRIOR FILING DATE: 1998-03-27
21	PRIOR APPLICATION NUMBER: 60/079786
22	PRIOR FILING DATE: 1998-03-27
23	PRIOR APPLICATION NUMBER: 60/079920
24	PRIOR FILING DATE: 1998-03-30
25	PRIOR APPLICATION NUMBER: 60/079923
26	PRIOR FILING DATE: 1998-03-30
27	PRIOR APPLICATION NUMBER: 60/080105
28	PRIOR FILING DATE: 1998-03-31
29	PRIOR APPLICATION NUMBER: 60/080107
30	PRIOR FILING DATE: 1998-03-31
31	PRIOR APPLICATION NUMBER: 60/080165
32	PRIOR FILING DATE: 1998-03-31
33	PRIOR APPLICATION NUMBER: 60/080194
34	PRIOR FILING DATE: 1998-03-31
35	PRIOR APPLICATION NUMBER: 60/080327
36	PRIOR FILING DATE: 1998-04-01
37	PRIOR APPLICATION NUMBER: 60/080328
38	PRIOR FILING DATE: 1998-04-01
39	PRIOR APPLICATION NUMBER: 60/080333
40	PRIOR FILING DATE: 1998-04-01
41	PRIOR APPLICATION NUMBER: 60/080334
42	PRIOR FILING DATE: 1998-04-01
43	PRIOR APPLICATION NUMBER: 60/081070
44	PRIOR FILING DATE: 1998-04-08
45	PRIOR APPLICATION NUMBER: 60/081049
46	PRIOR FILING DATE: 1998-04-08
47	PRIOR APPLICATION NUMBER: 60/081071
48	PRIOR FILING DATE: 1998-04-08
49	PRIOR APPLICATION NUMBER: 60/081195
50	PRIOR FILING DATE: 1998-04-08
51	PRIOR APPLICATION NUMBER: 60/081203
52	PRIOR FILING DATE: 1998-04-09
53	PRIOR APPLICATION NUMBER: 60/081229
54	PRIOR FILING DATE: 1998-04-09
55	PRIOR APPLICATION NUMBER: 60/081955
56	PRIOR FILING DATE: 1998-04-15
57	PRIOR APPLICATION NUMBER: 60/081817
58	PRIOR FILING DATE: 1998-04-15
59	PRIOR APPLICATION NUMBER: 60/081819
60	PRIOR FILING DATE: 1998-04-15
61	PRIOR APPLICATION NUMBER: 60/081952
62	PRIOR FILING DATE: 1998-04-15
63	PRIOR APPLICATION NUMBER: 60/081838
64	PRIOR FILING DATE: 1998-04-15
65	PRIOR APPLICATION NUMBER: 60/082568
66	PRIOR FILING DATE: 1998-04-21
67	PRIOR APPLICATION NUMBER: 60/082569
68	PRIOR FILING DATE: 1998-04-21
69	PRIOR APPLICATION NUMBER: 60/082704
70	PRIOR FILING DATE: 1998-04-22
71	PRIOR APPLICATION NUMBER: 60/082804
72	PRIOR FILING DATE: 1998-04-22
73	PRIOR APPLICATION NUMBER: 60/082700

PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22	PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23	PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27	PRIOR APPLICATION NUMBER: 60/083332
PRIOR FILING DATE: 1998-04-28	PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29	PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30	PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05	PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06	PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06	PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-5-07	PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07	PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13	PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13	PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13	PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15	PRIOR APPLICATION NUMBER: 60/085697

Query Match	100.0%;	Score 24;	DB 9;	Length 713;
Best Local Similarity	100.0%;	Pred. No. 5.6e+02;		

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLLVIA 6  
Db 454 LLLVIA 459

RESULT 9  
US-09-978-192A-183

; Sequence 183, Application US/09978192A  
; Patent No. US2002017553A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerltsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James:  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2630PIC9  
; CURRENT APPLICATION NUMBER: US/09/978,192A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 09/918585  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064249  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077641  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077791  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/078004  
; PRIOR FILING DATE: 1998-03-13  
; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078936  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078939

; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079664  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079663  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079786  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079920  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/079923  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/080105  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080107  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080165  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080194  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080327  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080328  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080333  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080334  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/081070  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081071  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081195  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081203  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081229  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081955  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081817  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081819  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081952  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081838  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082568  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082569  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082704  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082804  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082700  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082797  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082796  
; PRIOR FILING DATE: 1998-04-23  
; PRIOR APPLICATION NUMBER: 60/083366  
; PRIOR FILING DATE: 1998-04-27

PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

Query Match 100.0%; Score 24; DB 9; Length 713;  
Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LLLVTA 6  
Db 454 LLLVTA 459

RESULT 10  
US-09-999-832A-183  
Sequence 183, Application US/09999832A  
Publication No. US20020192706A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James:  
APPLICANT: Paonli, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumasz, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630PIC63  
CURRENT APPLICATION NUMBER: US/09/999,832A  
CURRENT FILING DATE: 2001-10-24  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496

PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

Query Match 100.0% Score 24; DB 9; Length 713;  
Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
DB 454 LLLVIA 459

RESULT 11  
US-10-052-586-416  
Sequence 416; Application US/10052586  
Patent No. US20020127584A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3430R1C1  
CURRENT APPLICATION NUMBER: US/10/052,586  
CURRENT FILING DATE: 2002-01-15  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059266  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/063120  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063121  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063486  
PRIOR FILING DATE: 1997-10-21  
PRIOR APPLICATION NUMBER: 60/063540  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063541  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063544  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063564  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063734  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063870  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/064103  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066120  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/066466  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066772  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/069335  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069425  
PRIOR FILING DATE: 1997-12-12  
PRIOR APPLICATION NUMBER: 60/069870  
PRIOR FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 60/068017  
PRIOR FILING DATE: 1997-12-18  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194

PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086023  
PRIOR FILING DATE: 1998-05-18  
PRIOR APPLICATION NUMBER: 60/086392  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/086486  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087098  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087208  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04

;; PRIOR APPLICATION NUMBER: 60/088033  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088167  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088202  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088212  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088217  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088326  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088655  
;; PRIOR FILING DATE: 1998-06-09  
;; PRIOR APPLICATION NUMBER: 60/088722  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088738  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088740  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088811  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088824  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088825  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088826  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088863  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089090  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089908

Query Match 100.0%; Score 24; DB 12; Length 713;  
Best Local Similarity 100.0%; Pred. No. 5.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LLLVIA 6  
Db 454 LLLVIA 459

RESULT 12  
US-09-529-063-70  
; Sequence 70, Application US/09529063  
; Patent No. US20020102542A1  
; GENERAL INFORMATION:  
; APPLICANT: FUKUSHIMA, DAICHI  
; APPLICANT: SHIBAYAMA, SHIRO  
; APPLICANT: TADA, HIDEAKI  
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF  
; FILE REFERENCE: Q58769  
; CURRENT APPLICATION NUMBER: US/09/529,063  
; CURRENT FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: PCT/JP98/04514

;; PRIOR FILING DATE: 1998-10-06  
;; PRIOR APPLICATION NUMBER: JP 9-274674  
;; PRIOR FILING DATE: 1997-10-07  
;; NUMBER OF SEQ ID NOS: 117  
;; SOFTWARE: Patentln Ver. 2.1  
;; SEQ ID NO 70  
;; LENGTH: 859  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-529-063-70

Query Match 100.0%; Score 24; DB 10; Length 859;  
Best Local Similarity 100.0%; Pred. No. 6.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LLLVIA 6  
Db 506 LLLVIA 511

RESULT 13  
US-09-984-245-179  
; Sequence 179, Application US/09984245  
; Patent No. US20020165374A1  
; GENERAL INFORMATION:  
; APPLICANT: Young et al.  
; TITLE OF INVENTION: 87 Human Secreted Proteins  
; FILE REFERENCE: P2004PI  
; CURRENT APPLICATION NUMBER: US/09/984,245  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 09/154,707  
; PRIOR FILING DATE: 1998-09-17  
; PRIOR APPLICATION NUMBER: PCT/US98/05311  
; PRIOR FILING DATE: 1998-03-19  
; PRIOR APPLICATION NUMBER: US 60/041,277  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/042,344  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,276  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/041,281  
; PRIOR FILING DATE: 1997-03-21  
; PRIOR APPLICATION NUMBER: US 60/048,094  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,350  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,188  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,135  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/050,937  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,187  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,099  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,352  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,186  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,069  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,095  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,131  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,096  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,355  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,160  
; PRIOR FILING DATE: 1997-05-30  
; PRIOR APPLICATION NUMBER: US 60/048,351



```

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 179
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-245-179

Query Match
Best Local Similarity 95.8%; Score 23; DB 9; Length 27;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6
   ||||:|
Db 7 LLLVIA 12

RESULT 14
US-09-864-761-34868
; Sequence 34868, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aomico-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
```

```

; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34868
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006977.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.3
; OTHER INFORMATION: SWISSPROT HIT: P00533, EVALUATE 9.00e-10
; OTHER INFORMATION: EST_HUMAN HIT: BE744284.1, EVALUATE 1.00e-08
US-09-864-761-34868

Query Match
Best Local Similarity 95.8%; Score 23; DB 10; Length 47;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6
   ||||:|
Db 16 LLLVIA 21

RESULT 15
US-10-001-870-116
; Sequence 116, Application US/10001870
; Patent No. US20020150924A1
; GENERAL INFORMATION:
; APPLICANT: Saiceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes a
; FILE REFERENCE: DEX-0283
; CURRENT APPLICATION NUMBER: US/10/001,870
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,189
; PRIOR FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 116
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-870-116

Query Match
Best Local Similarity 95.8%; Score 23; DB 12; Length 62;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

Search completed: January 3, 2003, 15:52:06  
Job time : 5.77273 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:21:45 : Search time 26.5909 Seconds  
(without alignments)  
30.067 Million cell updates/sec

Title: US-09-543-188a-13  
Perfect score: 24  
Sequence: 1 LLVIA 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A-Geneseq\_101002:\*  
1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	100.0	6	23	AAU11836
2	24	100.0	42	21	AAB34100
3	24	100.0	67	21	AAG6161
4	24	100.0	93	23	ABP05389
5	24	100.0	140	23	ABB90320
6	24	100.0	173	23	ABP39430
7	24	100.0	179	22	ABBI0539
8	24	100.0	179	22	AAU18139
9	24	100.0	179	22	AAU18690
10	24	100.0	179	22	AAU17055

11	24	100.0	182	21	AAG10813	Arabidopsis thalia
12	24	100.0	182	21	AAG42630	Arabidopsis thalia
13	24	100.0	183	21	AAG10812	Arabidopsis thalia
14	24	100.0	183	21	AAG42629	Arabidopsis thalia
15	24	100.0	271	22	AAU45545	Proionibacterium
16	24	100.0	283	22	ABG06298	Novel human diago
17	24	100.0	296	21	AAB42264	Human ORFX ORF2010
18	24	100.0	327	22	ABG82005	S. epidermidis ope
19	24	100.0	349	23	ABB91800	Herbivore-derived activ
20	24	100.0	366	22	ABB52990	Escherichia coli p
21	24	100.0	447	23	ABP28635	Streptococcus poly
22	24	100.0	478	22	AAU16984	Human novel secret
23	24	100.0	485	21	AAB59032	Breast and ovarian
24	24	100.0	487	22	ABBI0372	Human CDNA SEQ ID
25	24	100.0	491	23	ABB81636	Human zcytor19 pro
26	24	100.0	504	22	AAG90281	C glutamicum prote
27	24	100.0	520	23	ABB81643	Human zcytor19 pro
28	24	100.0	575	21	AAV71080	Murine TANGO 136 p
29	24	100.0	591	20	AAV15228	Human receptor pro
30	24	100.0	630	23	ABP39770	Staphylococcus epi
31	24	100.0	688	22	AAU93311	Human polypeptide,
32	24	100.0	713	20	AAV41712	Human PRO724 prote
33	24	100.0	713	21	AAB44268	Human PRO724 (UNQ3
34	24	100.0	713	21	AAV71081	Human TANGO 136 pr
35	24	100.0	713	22	AAU29231	Human PRO polypept
36	24	100.0	713	23	ABB95462	Human angiogenesis
37	24	100.0	713	23	ABB90346	Human polypeptide
38	24	100.0	713	23	ABB84856	Human PRO724 prote
39	24	100.0	713	23	ABB05751	Human G protein-co
40	24	100.0	814	22	ABG04441	Novel human diago
41	24	100.0	840	22	AAB70544	Human PRO14 protei
42	24	100.0	859	20	AAV02381	Polypeptide identi
43	24	100.0	859	21	AAB42317	Human ORFX ORF2081
44	24	100.0	859	22	AAU24052	Human EST encoded
45	24	100.0	859	22	AAU14316	Human novel protei

ALIGNMENTS

RESULT 1  
ID AAU11836 standard; peptide: 6 AA.  
AC AAU11836;  
DE 26-MAR-2002 (first entry)  
XX Peptide ligand for Prion protein, PrP. #11.  
XX  
KW Prion protein; PrP; ligand; octapeptide motif; scrapie;  
KW prion-associated disease; Creutzfeldt-Jakob disease;  
KW Gerstmann-Strausler-Scheinker disease; fatal familial insomnia;  
KW feline spongiform encephalopathy; bovine spongiform encephalopathy;  
KW transmissible mink encephalopathy; exotic ungulate encephalopathy;  
KW chronic wasting disease.  
XX  
OS Synthetic.  
XX  
PN WO200177687-A2.  
XX  
PD 18-OCT-2001.  
XX  
PF 05-APR-2001; 2001WO-US11150.  
XX  
PR 05-APR-2000; 2000US-0543188.  
XX  
PA (VITE-) VI TECHNOLOGIES INC.  
XX  
PI Hammond DJ, Wiltshire VR, Carbonell R, Shen H;  
XX  
DR WPI, 2002-061944/08.  
XX

PT New ligands for prion proteins, useful for detection or removal or  
PT prions and for treating prion-associated diseases, recognize a specific  
PT octapeptide motif -

XX Claim 16: Page 34; 47pp; English.

XX The invention relates to a ligand of less than 6 kD that binds to a  
CC polypeptide containing the sequence GLYTYRGLGYNProHISGLYGLY (A) or an  
CC analogue that is the retro-inverse isomer of (A). The sequence A is  
CC an octapeptide motif from the prion protein (PrP). The ligands are  
CC identified by binding assays with the peptide (A) or peptides containing  
CC (A). The ligands are used for detecting prion proteins (or prions) in  
CC biological or environmental samples, e.g. for diagnosis, also for  
CC removing them from samples, for treating or retarding development of  
CC prion-associated diseases, especially Creutzfeldt-Jakob diseases (in  
CC latent form, new variant, familial or sporadic forms), but also  
CC Gerstmann-Strausner-Scheinker disease, fatal familial insomnia,  
CC scrapie, bovine or feline spongiform encephalopathy, transmissible mink  
CC or exotic ungulate encephalopathy, or chronic wasting disease. The  
CC present sequence is a ligand of the invention.

Sequence 6 AA:

Query Match 100.0%; Score 24; DB 23; Length 6;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
DB 1 LLLVIA 6

RESULT 2  
AAB34100

ID AAB34100 standard; Protein: 42 AA.

AC AAB34100;

DT 26-JAN-2001 (first entry)

DE Human secreted protein sequence encoded by gene 9 SEQ ID NO:68.

XX Human; secreted protein; diagnosis; antiarthritic; immunosuppressive;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotrophic;  
KW cerebroprotective; neotrophic; neuroprotective; antibacterial; virucide;  
KW fungicide; ophthalmological; vulnary; gene therapy; autoimmune disease;  
KW hyperproliferative disorder; neoplasm; cancer; cardiovascular disorder;  
KW cerebrovascular disorder; angiogenesis; nervous system disorder;  
KW infection; ocular disorder; wound healing; skin aging; food additive;  
XX preservative.

OS Homo sapiens.

PN WO200056755-A1.

PD 28-SEP-2000.

PF 16-MAR-2000; 2000KO-US06830.

PR 19-MAR-1999; 99US-0125361.

PR 10-DEC-1999; 99US-0169910.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

DR WPI: 2000-587661/55.

DR N-PSDB: AAC59457.

XX New isolated nucleic acid molecules encoding 49 human secreted proteins  
PT used for preventing, treating or ameliorating medical conditions, for  
PT diagnosing pathological conditions or as food additives or  
PT preservatives

XX Claim 11: Page 366; 419pp; English.

XX The polynucleotide sequences given in AAC59449 to AAC59497 encode the  
CC human secreted proteins given in AAB34092 to AAB34140. AAB34141 to  
CC AAB34216 represent human secreted polypeptide sequences and proteins  
CC homologous to them, which are given in the exemplification of the present  
CC invention. Human secreted proteins have activities based on the tissue  
CC and cells the genes are expressed in. Examples of activities include:  
CC antirheumatic; immunosuppressive; antirheumatic; antiproliferative;  
CC cytostatic; cardiant; vasotrophic; cerebroprotective; neotrophic;  
CC neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
CC and vulnary. The polynucleotides and polypeptides can be used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Disorders which are diagnosed or treated include  
CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms or  
CC cancer of the breast or liver, cardiovascular disorders, cerebrovascular  
CC disorders, angiogenesis, nervous system disorders, infections caused by  
CC bacteria, viruses and fungi and ocular disorders. The polypeptides can  
CC also be used to aid wound healing and epithelial cell proliferation, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities. AAC59440 to AAC59448 and AAB34091 represent sequences used  
CC in the exemplification of the present invention.

Sequence 42 AA:

Query Match 100.0%; Score 24; DB 21; Length 42;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
DB 3 LLLVIA 8

RESULT 3  
AAG61661

ID AAG61661 standard; Protein: 67 AA.

AC AAG61661;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 80003.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149729.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151067.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0155658.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157733.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.



PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
PT prevention of neural, immune system, muscular, reproductive,  
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
PT disorders -  
XX Claim 11; SEQ ID NO 2696; 2081pp + Sequence Listing; English.  
XX  
CC The invention relates to novel genes (AB189449-AB190853) and proteins  
CC (AB889040-AB890444) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast; gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial infarction  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 140 AA:  
XX  
Query Match 100.0%; Score 24; DB 23; Length 140;  
Best Local Similarity 100.0%; Pred. No. 5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLLVIA 6  
Db 95 LLLVIA 100  
RESULT 6  
ABP39430  
ID ABP39430 standard; Protein: 173 AA.  
XX  
AC ABP39430;  
XX  
DT 24-JUL-2002 (first entry)  
XX  
DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4275.  
XX  
XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
XX antibacterial; gene therapy.  
XX  
OS Staphylococcus epidermidis.  
XX  
PN US6380370-B1.  
XX  
PD 30-APR-2002.  
XX  
PF 13-AUG-1998; 98US-0134001.  
XX  
PR 14-AUG-1997; 97US-055779P.  
XX  
PR 08-NOV-1997; 97US-064964P.  
XX  
PA (GENO-) GENOME THERAPEUTICS CORP.  
XX  
PI Doucette-Stamm LA, Bush D;  
XX  
DR WPI: 2002-381255/41.  
XX  
XX N-PSDB; ABN91975.  
PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis  
PT polypeptide, useful for diagnosing and treating bacterial infections -  
XX  
PS Disclosure: SEQ ID 4275; 267pp; English.  
XX

CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading  
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences  
CC given in ABP35124 to ABP37960. The S. epidermidis sequences have  
CC antibacterial activity and can be used in gene therapy. The sequences  
CC can also be used in the diagnosis and treatment of bacterial infections,  
CC particularly S. epidermidis infections. The sequences can be used to  
CC screen for compounds able to interfere with the S. epidermidis life  
CC cycle or inhibit S. epidermidis infection.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC USPTO web site.  
XX  
SQ Sequence 173 AA:  
XX  
Query Match 100.0%; Score 24; DB 23; Length 173;  
Best Local Similarity 100.0%; Pred. No. 6.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLLVIA 6  
Db 28 LLLVIA 33  
RESULT 7  
ABBI0539  
ID ABBI0539 standard; Protein: 179 AA.  
XX  
AC ABBI0539;  
XX  
DT 10-JAN-2002 (first entry)  
XX  
DE Human cDNA SEQ ID NO: 847.  
XX  
XX Human; gene therapy; neural disorder; immune system disorder;  
XX muscular disorder; reproductive disorder; gastrointestinal disorder;  
XX pulmonary disorder; cardiovascular disorder; renal disorder;  
XX proliferative disorder; inflammation.  
XX  
OS Homo sapiens.  
XX  
PN WO200154474-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001MO-US01349.  
XX  
XX 31-JAN-2000; 2000US-179065P.  
XX  
XX 04-FEB-2000; 2000US-180628P.  
XX  
XX 24-FEB-2000; 2000US-184664P.  
XX  
XX 02-MAR-2000; 2000US-186350P.  
XX  
XX 16-MAR-2000; 2000US-189874P.  
XX  
XX 17-MAR-2000; 2000US-190076P.  
XX  
XX 18-APR-2000; 2000US-198123P.  
XX  
XX 19-MAY-2000; 2000US-205515P.  
XX  
XX 07-JUN-2000; 2000US-209467P.  
XX  
XX 28-JUN-2000; 2000US-214886P.  
XX  
XX 30-JUN-2000; 2000US-215135P.  
XX  
XX 07-JUL-2000; 2000US-216647P.  
XX  
XX 07-JUL-2000; 2000US-216880P.  
XX  
XX 11-JUL-2000; 2000US-217487P.  
XX  
XX 11-JUL-2000; 2000US-217496P.  
XX  
XX 14-JUL-2000; 2000US-218290P.  
XX  
XX 26-JUL-2000; 2000US-220963P.  
XX  
XX 26-JUL-2000; 2000US-220964P.  
XX  
XX 14-AUG-2000; 2000US-224518P.  
XX  
XX 14-AUG-2000; 2000US-224519P.  
XX  
XX 14-AUG-2000; 2000US-225213P.  
XX  
XX 14-AUG-2000; 2000US-225214P.  
XX  
XX 14-AUG-2000; 2000US-225266P.  
XX  
XX 14-AUG-2000; 2000US-225267P.  
XX  
XX 14-AUG-2000; 2000US-225268P.  
XX  
XX 14-AUG-2000; 2000US-225270P.  
XX  
XX 14-AUG-2000; 2000US-225447P.  
XX

PR	14-AUG-2000	2000US-2255758
PR	14-AUG-2000	2000US-225758
PR	14-AUG-2000	2000US-225759
PR	18-AUG-2000	2000US-226279
PR	22-AUG-2000	2000US-226681
PR	22-AUG-2000	2000US-226688
PR	22-AUG-2000	2000US-227182
PR	23-AUG-2000	2000US-227000
PR	30-AUG-2000	2000US-228924
PR	01-SEP-2000	2000US-229287
PR	01-SEP-2000	2000US-229343
PR	01-SEP-2000	2000US-229344
PR	01-SEP-2000	2000US-229345
PR	05-SEP-2000	2000US-229509
PR	05-SEP-2000	2000US-229513
PR	05-SEP-2000	2000US-229538
PR	06-SEP-2000	2000US-230437
PR	06-SEP-2000	2000US-230438
PR	08-SEP-2000	2000US-231242
PR	08-SEP-2000	2000US-231243
PR	08-SEP-2000	2000US-231244
PR	08-SEP-2000	2000US-231413
PR	08-SEP-2000	2000US-231414
PR	08-SEP-2000	2000US-232068
PR	08-SEP-2000	2000US-232081
PR	12-SEP-2000	2000US-231968
PR	14-SEP-2000	2000US-232337
PR	14-SEP-2000	2000US-232338
PR	14-SEP-2000	2000US-232339
PR	14-SEP-2000	2000US-232400
PR	14-SEP-2000	2000US-232401
PR	14-SEP-2000	2000US-233063
PR	14-SEP-2000	2000US-233064
PR	14-SEP-2000	2000US-233065
PR	14-SEP-2000	2000US-233066
PR	21-SEP-2000	2000US-234233
PR	21-SEP-2000	2000US-234274
PR	25-SEP-2000	2000US-234997
PR	25-SEP-2000	2000US-235498
PR	26-SEP-2000	2000US-235484
PR	27-SEP-2000	2000US-235834
PR	27-SEP-2000	2000US-235836
PR	29-SEP-2000	2000US-236387
PR	29-SEP-2000	2000US-236388
PR	29-SEP-2000	2000US-236389
PR	29-SEP-2000	2000US-236390
PR	02-OCT-2000	2000US-237038
PR	02-OCT-2000	2000US-237039
PR	02-OCT-2000	2000US-237040
PR	13-OCT-2000	2000US-239395
PR	13-OCT-2000	2000US-239396
PR	20-OCT-2000	2000US-240967
PR	20-OCT-2000	2000US-241125
PR	20-OCT-2000	2000US-241126
PR	20-OCT-2000	2000US-241186
PR	20-OCT-2000	2000US-241187
PR	20-OCT-2000	2000US-241808
PR	20-OCT-2000	2000US-241809
PR	20-OCT-2000	2000US-241856
PR	01-NOV-2000	2000US-244617
PR	08-NOV-2000	2000US-246474
PR	08-NOV-2000	2000US-246475
PR	08-NOV-2000	2000US-246476
PR	08-NOV-2000	2000US-246477
PR	08-NOV-2000	2000US-246478
PR	08-NOV-2000	2000US-246534
PR	08-NOV-2000	2000US-246535
PR	08-NOV-2000	2000US-246556
PR	08-NOV-2000	2000US-246557
PR	08-NOV-2000	2000US-246578
PR	08-NOV-2000	2000US-246582

```

PR 08-NOV-2000; 2000US-246610P.
PR 08-NOV-2000; 2000US-246610P.
PR 08-NOV-2000; 2000US-246611P.
PR 08-NOV-2000; 2000US-246613P.
PR 17-NOV-2000; 2000US-249207P.
PR 17-NOV-2000; 2000US-249208P.
PR 17-NOV-2000; 2000US-249209P.
PR 17-NOV-2000; 2000US-249210P.
PR 17-NOV-2000; 2000US-249211P.
PR 17-NOV-2000; 2000US-249212P.
PR 17-NOV-2000; 2000US-249213P.
PR 17-NOV-2000; 2000US-249214P.
PR 17-NOV-2000; 2000US-249215P.
PR 17-NOV-2000; 2000US-249216P.
PR 17-NOV-2000; 2000US-249217P.
PR 17-NOV-2000; 2000US-249218P.
PR 17-NOV-2000; 2000US-249244P.
PR 17-NOV-2000; 2000US-249245P.
PR 17-NOV-2000; 2000US-249264P.
PR 17-NOV-2000; 2000US-249265P.
PR 17-NOV-2000; 2000US-249297P.
PR 17-NOV-2000; 2000US-249299P.
PR 17-NOV-2000; 2000US-249300P.
PR 01-DEC-2000; 2000US-250160P.
PR 01-DEC-2000; 2000US-250391P.
PR 05-DEC-2000; 2000US-251030P.
PR 05-DEC-2000; 2000US-251988P.
PR 05-DEC-2000; 2000US-256719P.
PR 06-DEC-2000; 2000US-251479P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
PR 08-DEC-2000; 2000US-251989P.
PR 08-DEC-2000; 2000US-251990P.
PR 11-DEC-2000; 2000US-254097P.
PR 05-JAN-2001; 2001US-259678P.

PA (HUMA-) HUMAN GENOME SCT INC.
XX
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI: 2001-476161/51.
DR N-PSDB; ABA06761.
XX
XX
XX Isolated nucleic acid molecule encoding an inflammation-associated
PT polypeptide is used in preventing, treating or ameliorating a medical
PI condition
XX
XX
XX
XX
XX Claim 11; SEQ ID NO: 847; 859pp + Sequence Listing; English.
PS
PS
XX
XX The present invention provides human CDNAs, proteins and related genomic
CC DNAs. These can be used in the treatment of neural, immune system,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal and proliferative disorders and inflammation. The present sequence
CC is a protein of the invention.
XX
XX
SQ Sequence 179 AA;

Query Match 100.0%; Score 24; DB 22; Length 179;
Best Local Similarity 100.0%; Pred. NO. 6.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY 1 LLLVIA 6
| | | | |
DB 134 LLLVIA 139

RESULT 8
AAU18139
ID AAU18139 standard; Protein; 179 AA.
XX
XX AAU18139;
XX

```



DT 21-NOV-2001 (first entry)  
XX Novel human uterine motility-association polypeptide #46.  
DE  
XX Human; uterine motility-association disorder; uterus; pregnancy;  
KW labour; menstrual cycle; gene therapy.  
XX  
XX Homo sapiens.  
OS  
PN WO200155201-A1.  
XX  
XX 02-AUG-2001.  
PD  
XX 17-JAN-2001; 2001WO-US01317.  
PF  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205315.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217495.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 14-AUG-2000; 2000US-0226279.  
PR 18-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236527.  
PR 29-SEP-2000; 2000US-0236567.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251785.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.



PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 05-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX  
DR WPI: 2001-48787/53.  
DR N-PSDB: AAS30211.  
XX  
XX  
PT New polynucleotides and polypeptides, useful for diagnosing, treating,  
PT preventing or prognosing e.g. kidney, cardiovascular, blood,  
PT electrolyte imbalance or neoplastic disorders, autoimmune diseases,  
PT cancers

XX  
PS Claim 1; SEQ ID NO 129; 506bp; English.  
XX  
CC The invention relates to novel nucleic acids and polypeptides useful for  
CC diagnosing, treating, preventing and/or prognosing disorders related to  
CC these polypeptides. The polynucleotides are especially useful in the  
CC diagnosis, prognosis, prevention and/or treatment of diseases which  
CC include kidney disorders (e.g. renal failure or nephritis),  
CC cardiovascular disorders (e.g. hypertension or myocardial infarction),  
CC blood disorders (e.g. anaemia or blood coagulation disorders),  
CC electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia),  
CC neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune  
CC diseases, cancers, inflammatory diseases, reproductive system  
CC disorders, endocrine disorders, neural activity and neurological  
CC disorders, wound healing and respiratory disorders. AAU1664-AAU18715  
CC represent the novel human renal and cardiovascular-associated amino  
CC acid sequences of the invention. Note: The sequence data for this patent  
CC did not form part of the printed specification, but was obtained in  
CC electronic format directly from WIPO at:  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 179 AA;  
Query Match 100.0%; Score 24; DB 22; Length 179;  
Best Local Similarity 100.0%; Pred. No. 6.6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LLLVIA 6  
|||||  
Db 134 LLLVIA 139  
RESULT 10  
AAU17055  
ID AAU17055 standard; Protein: 179 AA.  
XX  
AC AAU17055;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
XX  
DE Human novel secreted protein, SEQ ID 296.  
XX  
XX  
KW Human: immunosuppressive; antiarthritic; antirheumatic;  
KW cytosolic; cardiast; vasotropic; cerebroprotective; nootropic;  
KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
KW vulnary; secreted protein; rheumatoid arthritis;  
KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;  
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;  
KW corneal infection; wound healing; epithelial cell proliferation;  
KW skin ageing; food additive; preservative; antiproliferative.  
XX  
XX  
OS Homo sapiens.  
XX  
XX  
PN WO200155441-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01320.  
XX  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.

PR	11-JUL-2000;	2000US-0217487.
PR	11-JUL-2000;	2000US-0217496.
PR	11-JUL-2000;	2000US-0218290.
PR	26-JUL-2000;	2000US-0220963.
PR	26-JUL-2000;	2000US-0220964.
PR	14-AUG-2000;	2000US-0224518.
PR	14-AUG-2000;	2000US-0224519.
PR	14-AUG-2000;	2000US-0225211.
PR	14-AUG-2000;	2000US-0225214.
PR	14-AUG-2000;	2000US-0225266.
PR	14-AUG-2000;	2000US-0225267.
PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225447.
PR	14-AUG-2000;	2000US-0225757.
PR	14-AUG-2000;	2000US-0225758.
PR	14-AUG-2000;	2000US-0225759.
PR	18-AUG-2000;	2000US-0226279.
PR	22-AUG-2000;	2000US-0226681.
PR	22-AUG-2000;	2000US-0226686.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227009.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229509.
PR	05-SEP-2000;	2000US-0229513.
PR	06-SEP-2000;	2000US-0230437.
PR	06-SEP-2000;	2000US-0230438.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232080.
PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0231968.
PR	14-SEP-2000;	2000US-0232397.
PR	14-SEP-2000;	2000US-0232398.
PR	14-SEP-2000;	2000US-0232399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0233063.
PR	14-SEP-2000;	2000US-0233064.
PR	14-SEP-2000;	2000US-0233065.
PR	21-SEP-2000;	2000US-0234223.
PR	21-SEP-2000;	2000US-0234274.
PR	25-SEP-2000;	2000US-0234997.
PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235835.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239933.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	11-JUL-2000;	2000US-0217487.
PR	11-JUL-2000;	2000US-0217496.
PR	11-JUL-2000;	2000US-0218290.
PR	26-JUL-2000;	2000US-0220963.
PR	26-JUL-2000;	2000US-0220964.
PR	14-AUG-2000;	2000US-0224518.
PR	14-AUG-2000;	2000US-0224519.
PR	14-AUG-2000;	2000US-0225211.
PR	14-AUG-2000;	2000US-0225214.
PR	14-AUG-2000;	2000US-0225266.
PR	14-AUG-2000;	2000US-0225267.
PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225447.
PR	14-AUG-2000;	2000US-0225757.
PR	14-AUG-2000;	2000US-0225758.
PR	14-AUG-2000;	2000US-0225759.
PR	18-AUG-2000;	2000US-0226279.
PR	22-AUG-2000;	2000US-0226681.
PR	22-AUG-2000;	2000US-0226686.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227009.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;</	

CC immunosorbant assays (ELISA). Disorders which are diagnosed or treated  
CC include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi  
CC and ocular disorders e.g. corneal infection, and many other  
CC disorders listed in the specification. The polypeptides can also  
CC be used to aid wound healing and epithelial cell proliferation, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
CC minerals, cofactors and other nutritional components. The present

Query Match 100.0%; Score 24; DB 22; Length 179;  
Best Local Similarity 100.0%; Pred. No. 6,6e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LLLVIA 6  
134 LLLVIA 139

RESULT 11

ID AAG10813 standard; Protein; 162 AA.

XX AAG10813;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 9285.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0128845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 21-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.

```
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158363.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 24; DB 21; Length 182;
Best Local Similarity 100.0%; Pred. No. 6; +02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LLLVIA 6
Db 11 LLLVIA 16

RESULT 12
AAG42630
ID AAG42630 standard; Protein; 182 AA.
XX
AC AAG42630;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53187.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PM EPI033405-A2.
PD
XX
XX
PF 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
```

PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142820.  
PR 12-JUL-1999; 99US-0142877.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147792.  
PR 05-AUG-1999; 99US-0147760.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160880.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.

PR 29-OCT-1999: 99US-0162142.  
Query Match 100.0%; Score 24; DB 21; Length 182;  
Best Local Similarity 100.0%; Pred No. 6.7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLLVIA 6  
Db 11 LLLVIA 16  
RESULT 13  
AAG10812  
ID AAG10812 standard; Protein: 183 AA.  
XX AAG10812;  
AC  
XX  
DT 17-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 9284.  
Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.  
XX Arabidopsis thaliana.  
OS  
XX  
PN EPI033405-A2.  
PD  
XX 06-SEP-2000.  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999: 99US-0121825.  
PR 05-MAR-1999: 99US-0123180.  
PR 09-MAR-1999: 99US-0123548.  
PR 23-MAR-1999: 99US-0125788.  
PR 25-MAR-1999: 99US-0126264.  
PR 29-MAR-1999: 99US-0126785.  
PR 01-APR-1999: 99US-0127462.  
PR 06-APR-1999: 99US-0128234.  
PR 08-APR-1999: 99US-0128714.  
PR 16-APR-1999: 99US-0129845.  
PR 19-APR-1999: 99US-0130077.  
PR 21-APR-1999: 99US-0130449.  
PR 23-APR-1999: 99US-0130510.  
PR 28-APR-1999: 99US-0130891.  
PR 30-APR-1999: 99US-0131449.  
PR 30-APR-1999: 99US-0132048.  
PR 04-MAY-1999: 99US-0132407.  
PR 05-MAY-1999: 99US-0132484.  
PR 06-MAY-1999: 99US-0132485.  
PR 06-MAY-1999: 99US-0132486.  
PR 06-MAY-1999: 99US-0132487.  
PR 07-MAY-1999: 99US-0132863.  
PR 11-MAY-1999: 99US-0134256.  
PR 14-MAY-1999: 99US-0134218.  
PR 14-MAY-1999: 99US-0134219.  
PR 14-MAY-1999: 99US-0134221.  
PR 14-MAY-1999: 99US-0134370.  
PR 18-MAY-1999: 99US-0134768.  
PR 19-MAY-1999: 99US-0134768.  
PR 20-MAY-1999: 99US-0134941.  
PR 20-MAY-1999: 99US-0135124.  
PR 21-MAY-1999: 99US-0135533.  
PR 24-MAY-1999: 99US-0135629.  
PR 25-MAY-1999: 99US-0136021.  
PR 27-MAY-1999: 99US-0136392.  
PR 28-MAY-1999: 99US-0136782.  
PR 01-JUN-1999: 99US-0137222.  
PR 03-JUN-1999: 99US-0137528.  
PR 04-JUN-1999: 99US-0137502.  
PR 07-JUN-1999: 99US-0137724.  
PR 08-JUN-1999: 99US-0138094.

PR 10-JUN-1999: 99US-0138540.  
PR 10-JUN-1999: 99US-0138847.  
PR 14-JUN-1999: 99US-0139119.  
PR 16-JUN-1999: 99US-0139452.  
PR 16-JUN-1999: 99US-0139453.  
PR 17-JUN-1999: 99US-0139492.  
PR 18-JUN-1999: 99US-0139454.  
PR 18-JUN-1999: 99US-0139455.  
PR 18-JUN-1999: 99US-0139456.  
PR 18-JUN-1999: 99US-0139457.  
PR 18-JUN-1999: 99US-0139458.  
PR 18-JUN-1999: 99US-0139459.  
PR 18-JUN-1999: 99US-0139460.  
PR 18-JUN-1999: 99US-0139461.  
PR 18-JUN-1999: 99US-0139462.  
PR 18-JUN-1999: 99US-0139463.  
PR 18-JUN-1999: 99US-0139750.  
PR 18-JUN-1999: 99US-0139763.  
PR 21-JUN-1999: 99US-0139817.  
PR 22-JUN-1999: 99US-0139899.  
PR 23-JUN-1999: 99US-0140353.  
PR 23-JUN-1999: 99US-0140354.  
PR 24-JUN-1999: 99US-0140695.  
PR 28-JUN-1999: 99US-0140823.  
PR 29-JUN-1999: 99US-0140991.  
PR 30-JUN-1999: 99US-0141287.  
PR 01-JUL-1999: 99US-0141842.  
PR 01-JUL-1999: 99US-0142154.  
PR 02-JUL-1999: 99US-0142055.  
PR 06-JUL-1999: 99US-0142390.  
PR 08-JUL-1999: 99US-0142803.  
PR 09-JUL-1999: 99US-0142920.  
PR 12-JUL-1999: 99US-0142977.  
PR 13-JUL-1999: 99US-0143542.  
PR 14-JUL-1999: 99US-0143624.  
PR 15-JUL-1999: 99US-0144005.  
PR 16-JUL-1999: 99US-0144085.  
PR 16-JUL-1999: 99US-0144086.  
PR 19-JUL-1999: 99US-0144325.  
PR 19-JUL-1999: 99US-0144331.  
PR 19-JUL-1999: 99US-0144332.  
PR 19-JUL-1999: 99US-0144333.  
PR 19-JUL-1999: 99US-0144334.  
PR 19-JUL-1999: 99US-0144335.  
PR 20-JUL-1999: 99US-0144352.  
PR 20-JUL-1999: 99US-0144632.  
PR 20-JUL-1999: 99US-0144684.  
PR 21-JUL-1999: 99US-0144814.  
PR 21-JUL-1999: 99US-0145086.  
PR 21-JUL-1999: 99US-0145086.  
PR 22-JUL-1999: 99US-0145085.  
PR 22-JUL-1999: 99US-0145087.  
PR 22-JUL-1999: 99US-0145089.  
PR 22-JUL-1999: 99US-0145192.  
PR 23-JUL-1999: 99US-0145145.  
PR 23-JUL-1999: 99US-0145218.  
PR 23-JUL-1999: 99US-0145224.  
PR 26-JUL-1999: 99US-0145276.  
PR 27-JUL-1999: 99US-0145813.  
PR 27-JUL-1999: 99US-0145918.  
PR 27-JUL-1999: 99US-0145919.  
PR 28-JUL-1999: 99US-0145951.  
PR 02-AUG-1999: 99US-0146386.  
PR 02-AUG-1999: 99US-0146388.  
PR 02-AUG-1999: 99US-0146389.  
PR 03-AUG-1999: 99US-0147038.  
PR 04-AUG-1999: 99US-0147204.  
PR 04-AUG-1999: 99US-0147302.  
PR 05-AUG-1999: 99US-0147192.  
PR 05-AUG-1999: 99US-0147260.  
PR 06-AUG-1999: 99US-0147303.  
PR 06-AUG-1999: 99US-0147416.  
PR 09-AUG-1999: 99US-0147493.



PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157863.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0158294.  
PR 13-OCT-1999; 99US-0158295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 24; DB 21; Length 183;  
Best Local Similarity 100.0%; Pred. No. 6; Be+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 12 LTLVIA 17  
RESULT 14  
AAG42629  
ID AAG42629 standard; Protein; 183 AA.  
XX  
AC AAG42629;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 53186.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.

```
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139765.
PR 18-JUN-1999; 99US-0139766.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140822.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145986.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155658.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157755.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match          100.0%; Score 24; DB 21; Length 183;
Best Local Similarity 100.0%; Pred. No. 6; Be+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 LLLVIA 6
Db      12 LLLVIA 17
```

```
RESULT 15
AA045545
ID AA045545 standard; Protein; 271 AA.
```

XX AAU45545;  
AC  
XX  
DT 27-FEB-2002 (first entry)  
XX  
DE Propionibacterium acnes immunogenic protein #6441.  
XX  
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
OS Propionibacterium acnes.  
XX  
XX WO200181581-A2.  
XX  
XX 01-NOV-2001.  
XX  
XX 20-APR-2001; 2001WO-US12865.  
XX  
XX 21-APR-2000; 2000US-199047P.  
XX 02-JUN-2000; 2000US-208641P.  
XX 07-JUL-2000; 2000US-216747P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX WPI: 2001-616774/71.  
XX N-PSDB; AAS59526.  
XX  
XX Propionibacterium acnes polypeptides and nucleic acids useful for  
XX PT vaccinating against and diagnosing infections, especially useful for  
XX PT treating acne vulgaris.  
XX  
XX Claim 3: SEQ ID NO 6740; 1069pp; English.  
XX  
XX Sequences AAU939105-AAU68017 represent Propionibacterium acnes immunogenic  
XX polypeptides. The proteins and their associated DNA sequences are used in  
XX the treatment, prevention and diagnosis of medical conditions caused by  
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
XX P. acnes is also involved in infections of bone, joints and the central  
XX nervous system, however it is particularly involved in the inflammatory  
XX lesions associated with acne vulgaris. A method for detecting the  
XX presence or absence of P. acnes in a patient comprises contacting a  
XX sample with a binding agent that binds to the proteins of the invention  
XX and determining the amount of bound protein in the sample. The  
XX polypeptides may be used as antigens in the production of antibodies  
XX specific for P. acnes proteins. These antibodies can be used to  
XX downregulate expression and activity of P. acnes polypeptides and  
XX therefore treat P. acnes infections. The antibodies may also be used as  
XX diagnostic agents for determining P. acnes presence, for example, by  
XX enzyme linked immunosorbent assay (ELISA).  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
XX SQ Sequence 271 AA;

Query Match 100.0%; Score 24; DB 22; Length 271;  
Best Local Similarity: 100.0%; Pred. No. 1e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLLVIA 6  
| | | | |  
Db 47 LLLVIA 52

Search completed: January 3, 2003, 15:28:38  
Job time : 27.5909 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:00 ; Search time 9.95455 Seconds  
(without alignments)  
57.944 Million cell updates/sec

Title: US-09-543-188a-13

Perfect score: 24

Sequence: 1 LLLVIA 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

al number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	100.0	150	2 C71515	probable 15kda cys
2	24	100.0	152	2 F43584	15k cysteine-rich
3	24	100.0	159	2 D82812	hypothetical prote
4	24	100.0	183	2 G84609	hypothetical prote
5	24	100.0	224	2 T32185	hypothetical prote
6	24	100.0	232	2 C72609	hypothetical prote
7	24	100.0	239	2 T36065	probable integral
8	24	100.0	276	1 PMYBA	H+-transporting tw
9	24	100.0	277	2 T29611	hypothetical prote
10	24	100.0	281	2 S74551	hypothetical prote
11	24	100.0	284	2 S22414	signal peptidase I
12	24	100.0	284	2 H83550	probable signal pe
13	24	100.0	295	2 G98301	hypothetical ABC t
14	24	100.0	295	2 A12881	ABC transporter, m
15	24	100.0	295	2 AD0306	probable membrane
16	24	100.0	296	2 T12469	hypothetical prote
17	24	100.0	297	2 S75656	hypothetical prote
18	24	100.0	301	2 H97505	probable nKc prot
19	24	100.0	301	2 AD2724	ABC transporter, m
20	24	100.0	316	2 T31680	hypothetical prote
21	24	100.0	326	2 B71644	undecaprenyl-phosp
22	24	100.0	349	2 T00525	probable GDSL-moti
23	24	100.0	361	2 G72636	hypothetical prote
24	24	100.0	375	2 H86938	conserved hypotet
25	24	100.0	392	2 D70475	conserved hypotet
26	24	100.0	428	2 D70583	probable transmemb
27	24	100.0	437	2 S73284	probable transmemb
28	24	100.0	464	2 E64060	p1b protein homol
29	24	100.0	487	2 E70480	probable sodium/pr

30	24	100.0	509	2 A57512	beta-glucosidase B
31	24	100.0	601	2 E95296	probable ABC-type
32	24	100.0	654	2 G84153	hypothetical prote
33	24	100.0	726	2 S50833	holocarboxylase sy
34	24	100.0	770	2 T00204	LDL receptor relat
35	24	100.0	770	2 T00203	LDL receptor relat
36	24	100.0	896	2 AB1156	conserved membrane
37	24	100.0	896	2 AE1514	conserved membrane
38	24	100.0	1011	2 T45718	receptor-kinase 11
39	24	100.0	1239	2 T42020	class IV chitin sy
40	24	100.0	1266	2 T22090	hypothetical prote
41	23	95.8	50	2 D72804	gp38 protein - Myc
42	23	95.8	73	2 AB3444	hypothetical prote
43	23	95.8	100	2 T43116	hypothetical prote
44	23	95.8	104	2 AC0162	probable membrane
45	23	95.8	105	2 F70545	hypothetical prote

## ALIGNMENTS

## RESULT 1

C71515  
probable 15kda cysteine-rich protein - Chlamydia trachomatis (serotype D, strain UW3/  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 10-Dec-1999  
C:Accession: C71515  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Martche, R.; Aravind, L.; Mitche  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t  
A:Reference number: A71570; WUID:9900809; PMID:9784136  
A:Accession: C71515  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-150 <ARN>  
A:Cross-references: GB:AE001317; GB:AE001273; MID:93328663; PIDN:AAC68041.1; PID:9332  
A:Experimental source: serotype D, strain UW-3/CX  
C:Genetics:  
A:Gene: crpA  
C:Superfamily: Chlamydia 15k cysteine-rich outer membrane protein

Query Match 100.0%; Score 24; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6

DB 48 LLLVIA 53

## RESULT 2

F43584  
15k cysteine-rich outer membrane protein - Chlamydia trachomatis (serotype C)  
C:Species: Chlamydia trachomatis  
C:Date: 12-Jan-1993 #sequence\_revision 12-Jan-1993 #text\_change 10-Dec-1999  
C:Accession: F43584; S24274  
R:de la Maza, L.M.; Fielder, T.J.; Carlson, E.J.; Marroff, B.A.; Peterson, E.M.  
Infect. Immun. 59, 1196-1201, 1991  
A:Title: Sequence diversity of the 60-kilodalton protein and of a putative 15-kilodal  
A:Reference number: A43584; WUID:91147205; PMID:1997423  
A:Accession: F43584  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-152 <DEL>  
A:Cross-references: EMBL:X54388; NID:940760; PIDN:CAA38258.1; PID:940762  
C:Superfamily: Chlamydia 15k cysteine-rich outer membrane protein  
C:Keywords: membrane protein

Query Match 100.0%; Score 24; DB 2; Length 152;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLLVIA 6



A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-239 <SAU>  
A:Cross-references: EMBL:AL034355; PIDN:CAA22231.1; GSPDB:GN00070; SCOPDB:SCD78.26  
C:Genetics:  
A:Gene: SCOPDB:SCD78.26

Query Match  
Best Local Similarity 100.0%; Score 24; DB 2; Length 239;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||  
DB 29 LLLVIA 34

## RESULT 8

PHYBAA  
transporting two-sector ATPase (EC 3.6.3.14) chain a - Synechocystis sp. (strain PCC  
Alternate names: ATP synthase chain a  
Species: Synechocystis sp.

A:Variety: PCC 6803  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 14-Dec-2001  
C:Accession: S17746; S74588; S14861  
R:Liill, H.; Nelson, N.  
Plant Mol. Biol. 17, 641-652, 1991  
A:Title: The atp1 and atp2 operons of the cyanobacterium Synechocystis sp. PCC 6803.  
A:Reference number: S17745; MUID:92003679; PMID:1832989

A:Accession: S17746  
A:Molecule type: DNA  
A:Residues: 1-276 <LII>  
A:Cross-references: EMBL:X58128; NID:g47506; PIDN:CAA41130.1; PID:g47508  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
s.  
A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S74588  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-276 <KAN>  
A:Cross-references: EMBL:D90900; GB:AB001339; NID:g1651768; PIDN:BA16740.1; PID:g165181  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:

A:Gene: atp1  
A:Superfamily: H+-transporting ATP synthase protein 6  
A:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; transmembrane prot  
F:63-82/Domain: transmembrane #status predicted <TM1>  
F:122-140/Domain: transmembrane #status predicted <TM2>  
F:160-178/Domain: transmembrane #status predicted <TM3>  
F:223-242/Domain: transmembrane #status predicted <TM4>  
F:247-265/Domain: transmembrane #status predicted <TM6>

Query Match  
Best Local Similarity 100.0%; Score 24; DB 1; Length 276;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||  
DB 72 LLLVIA 77

## RESULT 9

T29611  
hypothetical protein C54H2.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000  
C:Accession: T29611  
R:Fulton, L.; Gattung, S.  
submitted to the EMBL Data Library, May 1996

A:Description: The sequence of C. elegans cosmid C54H2.  
A:Reference number: 220651  
A:Accession: T29611

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-277 <FNU>  
A:Cross-references: EMBL:U58728; PIDN:AAB00591.1; GSPDB:GN00028; CESP:C54H2.5  
A:Experimental source: strain Bristol N2; clone C54H2  
C:Genetics:

A:Gene: CESP:C54H2.5  
A:Map position: X  
A:Introns: 23/3; 126/2; 230/2  
C:Superfamily: Caenorhabditis elegans hypothetical protein C54H2.5

Query Match  
Best Local Similarity 100.0%; Score 24; DB 2; Length 277;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||  
DB 257 LLLVIA 262

## RESULT 10

S74551  
hypothetical protein slr0241 - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.

A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S74551  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
s.

A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S74551  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-281 <KAN>  
A:Cross-references: EMBL:D90900; GB:AB001339; NID:g1651768; PIDN:BA16703.1; PID:d101  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match  
Best Local Similarity 100.0%; Score 24; DB 2; Length 281;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||  
DB 264 LLLVIA 269

## RESULT 11

S22414  
signal peptidase I (EC 3.4.21.89) - Pseudomonas fluorescens  
N:Alternate names: leader peptidase 1  
C:Species: Pseudomonas fluorescens  
C:Date: 22-Nov-1993 #sequence\_revision 01-Dec-1995 #text\_change 19-May-2000  
C:Accession: S22414  
R:Black, M.T.; Munn, J.G.R.; Allsop, A.E.  
Biochem. J. 282, 539-543, 1992

A:Title: On the catalytic mechanism of prokaryotic leader peptidase 1.  
A:Reference number: S22413; MUID:92189595; PMID:1546569

A:Accession: S22414  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-284 <BLA>  
A:Cross-references: EMBL:X56466; NID:g45509; PIDN:CAA39839.1; PID:g45511  
C:Superfamily: signal peptidase I  
C:Keywords: hydrolase; serine proteinase; transmembrane protein

Query Match  
Best Local Similarity 100.0%; Score 24; DB 2; Length 284;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||

DB 7 LLLVIA 12

## RESULT 12

H83550

probable signal peptidase I (EC 3.4.21.89) PA0768 [similarity] - *Pseudomonas aeruginosa*  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 24-May-2001

C:Accession: H83550  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
A:Reference number: AB2950; MUID:20437337; PMID:10984043  
C:Accession: H83550  
Status: preliminary

Molecule type: DNA

A:Residues: 1-284 &lt;STO&gt;

A:Cross-references: GB:AE004511; GB:AE004091; NID:g9946646; PIDN:AAG04157.1; GSPDB:GN001  
A:Experimental source: strain PA01

C:Genetics:

A:Gene: lepb; PA0768

C:Superfamily: signal peptidase I

C:Keywords: hydrolase; serine proteinase

Query Match 100.0%; Score 24; DB 2; Length 284;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||

DB 7 LLLVIA 12

## RESULT 13

G98301

hypothetical ABC transporter permease protein yddQ AGR\_L\_2741 [imported] - *Agrobacterium*  
C:Species: *Agrobacterium tumefaciens*  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002

C:Accession: G98301

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Mollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*  
A:Reference number: A97359; PMID:11743194

A:Accession: G98301

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-295 &lt;KUR&gt;

A:Cross-references: GB:AE007870; PIDN:AAK89937.1; PID:g15159892; GSPDB:GN00170

C:Genetics:

A:Gene: AGR\_L\_2741

A:Map position: linear chromosome

C:Superfamily: oligopeptide permease protein oppb

Query Match 100.0%; Score 24; DB 2; Length 295;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||

DB 39 LLLVIA 44

## RESULT 14

A12981

ABC transporter, membrane spanning protein Atu3458 [imported] - *Agrobacterium tumefaciens*  
C:Species: *Agrobacterium tumefaciens*



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:22:10 : Search time 5.18182 Seconds  
(without alignments)  
48.025 Million cell updates/sec

Title: US-09-543-188a-13  
Perfect score: 24  
Sequence: 1 LLLVIA 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
al number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	100.0	96	BV8_BOMVA	Q9PW66 bombina var
2	24	100.0	117	Y13K_BPT4	P39504 bacterioph
3	24	100.0	150	CRPD_CHLTR	O84449 chlamydia t
4	24	100.0	152	CRPC_CHLTR	P26756 chlamydia t
5	24	100.0	162	Y066_GVCL	P41728 cryptophleb
6	24	100.0	276	ATP6_STNY3	P27178 synecocyst
7	24	100.0	277	SUR4_CAEL	Q18864 caenorhabd
8	24	100.0	284	LEP_PSEAE	Q91597 pseudomonas
9	24	100.0	284	LEP_PSEEL	P26844 pseudomonas
10	24	100.0	415	MNTH_MYCBO	O69443 mycobacteri
11	24	100.0	428	MNTH_MYCTU	O05916 mycobacteri
12	24	100.0	437	YC44_PORPU	P51363 porphyra pu
13	24	100.0	464	HOPE_HAEIN	P44622 haemophilus
14	24	100.0	726	BPL1_HUMAN	P50747 h biotin--P
15	24	100.0	1239	CHS5_USTMA	O13394 ustilago ma
16	24	95.8	50	VG38_BRMD2	O64229 mycobacteri
17	23	95.8	145	Y657_ARCFU	O29600 archaeoglob
18	23	95.8	188	DSBE_YERPE	O82d52 yersinia pe
19	23	95.8	205	Y115_NPVOP	O10354 oryza psen
20	23	95.8	207	IL6_MARMO	O35736 marmota mon
21	23	95.8	208	IL6_FELCA	P41683 felis silve
22	23	95.8	246	VS09_ROTGI	P30889 rotavirus (
23	23	95.8	269	PTFC_BACSU	P26381 bacillus su
24	23	95.8	269	SUR4_FUGRU	O57580 fugu rubrip
25	23	95.8	269	SUR4_HUMAN	O15260 homo sapien
26	23	95.8	269	SUR4_MOUSE	O64310 mus musculi
27	23	95.8	318	FECD_ECOLI	P15029 escherichia
28	23	95.8	334	FECD_ECOLI	P23876 escherichia
29	23	95.8	363	CYA2_RHIME	O52915 rhizobium m
30	23	95.8	425	GSPU_ERMCA	P31708 aquifex car
31	23	95.8	429	SECY_AQUAE	O66491 aquifex aeo
32	23	95.8	430	YBAT_ECOLI	P77400 escherichia
33	23	95.8	499	C831_ARATH	O65782 arabidopsis

34	23	95.8	503	1	AGP4_HUMAN	O9Y264 homo sapien
35	23	95.8	531	1	PTM1_YEAST	P32857 saccharomyc
36	23	95.8	628	1	LU_HUMAN	P50895 homo sapien
37	23	95.8	633	1	NUOL_MYCTU	O86350 mycobacteri
38	23	95.8	704	1	PMA2_LYCES	P23960 lycopersico
39	23	95.8	910	1	DDRI_RAT	O63474 rattus norv
40	23	95.8	911	1	DDRI_MOUSE	O03146 mus musculu
41	23	95.8	913	1	DDRI_HUMAN	O08345 h epithelia
42	23	95.8	956	1	PMA3_NICPL	O08436 nicotiana p
43	23	95.8	982	1	YS96_CAEL	O09965 caenorhabd
44	23	95.8	1047	1	ANPB_BOVIN	P46197 bos taurus
45	23	95.8	1047	1	ANPB_RAT	P16067 rattus norv

## ALIGNMENTS

RESULT 1	ID	STANDARD	PRT	96 AA
BV8_BOMVA				
AC	Q9PW66;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Protein BV8 precursor.			
OS	Bombina variegata (Yellow-bellied toad).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Archaeobatrachia; Bombinatoridae; Bombina.			
OX	NCBI_TaxID=8348;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	TISSUE=SKIN.			
RX	MEDLINE=99349621; PubMed=10422759;			
RA	Mollay C., Wechseltberger C., Mignogna G., Negri L., Melchiorri P.,			
RA	Barra D., Kreil G.;			
RT	"BV8, a small protein from frog skin and its homologue from snake			
RT	venom induce hyperalgesia in rats.";			
RL	Eur. J. Pharmacol. 374:189-196(1999).			
CC	-!- FUNCTION: Potently contract gastrointestinal (GI) smooth muscle.			
CC	-!- INDUCES hyperalgesia.			
CC	-!- SUBCELLULAR LOCATION: Secreted.			
CC	-!- SIMILARITY: BELONGS TO THE PROKINECITIN FAMILY.			
CC	-!- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
DR	EMBL: AF168790; AAD45816.1; -.			
DR	HSSP: P25687; 1IMT.			
KW	Signal.			
FT	CHAIN 1 19			
FT	DISULFID 26 38			
FT	DISULFID 32 50			
FT	DISULFID 37 78			
FT	DISULFID 60 86			
FT	DISULFID 80 95			
FT	SEQUENCE 96 AA; 10102 MW; A12490A7437609B4 CRC64;			
QY	1 LLLVIA 6			
QY				
QY	10 LLLVIA 15			
DB	10 LLLVIA 15			
RESULT 2				

Query Match 100.0%; Score 24; DB 1; Length 96;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 6; Conservative 0; Mismatches 0;  
Indels 0; Gaps 0;

Y13K\_BP74  
ID Y13K\_BP74 STANDARD; PRT; 117 AA.  
AC P39504;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Hypothetical 13.1 kDa protein in pset-alc intergenic region.  
GN Y13K OR PSET.3.  
OS Bacteriophage T4.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;  
OC T4-like viruses.  
OC NCBI\_TaxID=10665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,  
RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;  
RT "Bacteriophage T4 genome analysis."  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: AF158101; AAD42645.1; -  
CC DR Hypothetical protein.  
CC KM  
CC SQ SEQUENCE 117 AA; 13135 MW; D8B52577D7DA979B CRC64;

Query Match 100.0%; Score 24; DB 1; Length 117;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLVIA 6  
DB 11 LLVIA 16

RESULT 3  
CRPD\_CHLTR  
ID CRPD\_CHLTR STANDARD; PRT; 150 AA.  
AC 084449;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 15 kDa cysteine-rich outer membrane protein, serovar D.  
GN CRPA OR CT442.  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OC NCBI\_TaxID=813;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=D/UN-3/CX;  
RA MEDLINE=9000809; PubMed=9784136;  
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
RA Davis R.W.;  
RT "Genome sequence of an obligate intracellular pathogen of humans:  
RT Chlamydia trachomatis."  
RL Science 282:754-759(1998).  
CC -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES  
CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL  
CC INTEGRITY OF THE EBS OUTER ENVELOPE.  
CC -1- SUBCELLULAR LOCATION: Outer membrane.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AE001317; AAC68041.1; -  
CC DR Outer membrane; Complete proteome.  
CC KM  
CC SQ SEQUENCE 150 AA; 15985 MW; 77A6FB7CC8FAD34 CRC64;

Query Match 100.0%; Score 24; DB 1; Length 150;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLVIA 6  
DB 48 LLVIA 53

RESULT 4  
CRPD\_CHLTR  
ID CRPD\_CHLTR STANDARD; PRT; 152 AA.  
AC P26756;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE 15 kDa cysteine-rich outer membrane protein, serovar C.  
GN CRPA.  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OC NCBI\_TaxID=813;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=C/TW-3;  
RA MEDLINE=91147205; PubMed=1997423;  
RA de la Maza L.M., Fiedler T.J., Carlson E.J., Markoff B.A.,  
RA Peterson E.M.;  
RT "Sequence diversity of the 60-kilodalton protein and of a putative  
RT 15-kilodalton protein between the trachoma and lymphogranuloma  
RT venereum biovars of Chlamydia trachomatis."  
RL Infect. Immun. 59:1196-1201(1991).  
CC -1- FUNCTION: ASSOCIATED WITH DIFFERENTIATION OF RETICULATE BODIES  
CC (RBS) INTO ELEMENTARY BODIES (EBs). IT IS NECESSARY FOR STRUCTURAL  
CC INTEGRITY OF THE EBS OUTER ENVELOPE.  
CC -1- SUBCELLULAR LOCATION: Outer membrane.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: X54388; CAA38258.1; -  
CC DR PIR: F43584; F43584.  
CC KM Outer membrane.  
CC SQ SEQUENCE 152 AA; 16133 MW; 31D325CC0619A2F1 CRC64;

Query Match 100.0%; Score 24; DB 1; Length 152;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLVIA 6  
DB 50 LLVIA 55

RESULT 5  
Y096\_GVCL  
ID Y096\_GVCL STANDARD; PRT; 162 AA.  
AC P41728;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Hypothetical 18.6 kDa protein in P143-LfE5 intergenic region.  
OS Cryptophlebia leucotreta granulosis virus (ClGV) (Cryptophlebia

OS leucotreta granulovirus).  
OC VIRUSES: dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.  
OX NCBI\_TaxID=35254;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94292936; PubMed=8021613;  
RA Jehle J.A., Backhaus H.;  
RT "Genome organization of the DNA-binding protein gene region of  
RT Cryptophlebia leucotreta granulosis virus is closely related to that  
RT of nuclear polyhedrosis viruses";  
RL J. Gen. Virol. 75:1815-1820(1994).  
CC -1- SIMILARITY: TO CORRESPONDING ORF IN ACMPV AND OPMNV.  
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X77048; CAA54340.1; -.  
CC KW Hypothetical protein.  
SQ SEQUENCE 162 AA; 18556 MW; 97C75FD0DE9FP88 CRC64;  
Query Match 100.0%; Score 24; DB 1; Length 162;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLLVIA 6  
Db 9 LLLVIA 14  
RESULT 6  
ATP6\_SYNV3  
ID ATP6\_SYNV3 STANDARD; PRT; 276 AA.  
AC P27178;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE ATP synthase A chain (EC 3.6.3.14) (Protein 6).  
GN ATPB OR ATP1 OR SLI1322.  
OS Synchocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.  
OX NCBI\_TaxID=1148;  
[1]  
SEQUENCE FROM N.A.  
RA Lill H., Nelson N.; PubMed=1832989;  
RT "The atp1 and atp2 operons of the cyanobacterium Synchocystis sp.  
RT PCC 6803.";  
RL Plant Mol. Biol. 17:641-652(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Sugiyura M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,  
RA Okumura S., Shimo S., Takeuchi C., Wada T., Watanabe A.,  
RA Yamada M., Yasuda M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synchocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A  
CC DIRECT ROLE IN THE TRANSLLOCATION OF PROTONS ACROSS THE MEMBRANE.  
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +  
CC H(+) (out).  
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC  
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE  
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)

CC HAS THREE MAIN SUBUNITS: A, B AND C.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. CONTAINS 8  
CC POTENTIAL TRANSMEMBRANE DOMAINS.  
CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X58128; CAA41130.1; -.  
CC DR EMBL: D90900; BAA16740.1; -.  
CC DR PIR: S17746; PWBDA.  
CC DR InterPro: IPR000568; ATPsyn\_Asub.  
CC DR Pfam: PF00119; ATP-synt\_A; 1.  
CC DR PRINTS: PR00123; ATPASEA.  
CC DR TIGRFS: TIGR01131; ATP-synt\_6\_or\_A; 1.  
CC DR PROSITE: PS00449; ATPASE\_A; 1.  
CC KW Hydrogen ion transport; CF(0); Transmembrane; Complete proteome.  
FT TRANSMEM 27 44 POTENTIAL.  
FT TRANSMEM 63 82 POTENTIAL.  
FT TRANSMEM 122 140 POTENTIAL.  
FT TRANSMEM 160 179 POTENTIAL.  
FT TRANSMEM 225 242 POTENTIAL.  
FT TRANSMEM 247 265 POTENTIAL.  
SQ SEQUENCE 276 AA; 30698 MW; E65B9DE31C7621FF CRC64;  
Query Match 100.0%; Score 24; DB 1; Length 276;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLLVIA 6  
Db 72 LLLVIA 77  
RESULT 7  
SUR4\_CAEEL  
ID SUR4\_CAEEL STANDARD; PRT; 277 AA.  
AC Q18864; Q18006;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Surfeit locus protein 4 homolog.  
GN SURF-4 OR C54H2.5.  
OS Caenorhabditis elegans.  
OC Caenorhabditis elegans.  
OC Rhadylidae; Metazoa; Nematoda; Chromadorea; Rhadylidae; Rhadylidae;  
OX NCBI\_TaxID=6239;  
[1]  
SEQUENCE FROM N.A.  
RX MEDLINE=96413309; PubMed=8816471;  
RA Ames N., Fried M.;  
RT "Surfeit locus gene homologs are widely distributed in invertebrate  
RT genomes.";  
RL Mol. Cell. Biol. 16:5591-5596(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX SRRATN-Bristol N2;  
RA Fulton L., Gattung S.;  
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic  
CC reticulum (by similarity).  
CC -1- SIMILARITY: BELONGS TO THE SURF4 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

```

CC entties requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; U58728; AAB00591.1; -.
CC DR EMBL; Y14949; CAA75173.1; -.
CC DR Wormpep; C54H2.5; CE06987.
CC DR InterPro; IPR002995; Surf4.
CC DR Pfam; PF02077; SURF4; 1.
CC DR ProDom; PD010195; Surf4; 1.
CC DR PROSITE; PS01339; SURF4; 1.
CC KW Transmembrane; Endoplasmic reticulum.
CC FT TRANSMEM 71 91 POTENTIAL.
CC FT TRANSMEM 97 117 POTENTIAL.
CC FT TRANSMEM 118 138 POTENTIAL.
CC FT TRANSMEM 163 183 POTENTIAL.
CC FT TRANSMEM 187 207 POTENTIAL.
CC FT TRANSMEM 213 233 POTENTIAL.
CC FT TRANSMEM 250 270 POTENTIAL.
CC FT SITE 274 275 POTENTIAL.
CC -----
CC KW ENDOPLASMIC RETICULUM RETRIEVAL MOTIF
CC (POTENTIAL).
CC A -> R (IN REF. 1).
CC SEQUENCE 277 AA; 31813 MW; 8FF2848676CC0245 CRC64;
CC -----
CC Query Match 100.0%; Score 24; DB 1; Length 277;
CC Best Local Similarity 100.0%; Pred. No. 2e+02;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 1 LLLVIA 6
CC DB 257 LLLVIA 262
CC -----
CC RESULT 8
CC LEPP_PSEAE
CC ID LEPP_PSEAE STANDARD; PRT; 284 AA.
CC AC 0915G7;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Signal peptidase I (EC 3.4.21.89) (Spase I) (Leader peptidase I).
CC GN LEPP OR P40768.
CC OS Pseudomonas aeruginosa.
CC OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
CC OC Pseudomonas.
CC OX NCBI_TaxID=287;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=ATCC 15692 / PAOI;
CC MEDLINE=20437337; Pubmed=10984043;
CC RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
CC RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
CC RA Gardner R.L., Goltzy L., Tolentino E., Westbrock-Adams S., Yuan Y.,
CC RA Bader L.L., Coulter S.N., Folger K.R., Kas A., Labdig K., Lim R.M.,
CC RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
CC RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
CC RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
CC RT opportunistic pathogen."
CC RL Nature 406:959-964(2000).
CC CC -I- CATABOLIC ACTIVITY: Cleavage of N-terminal leader sequences from
CC CC secreted and periplasmic proteins precursor.
CC CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
CC CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S26; ALSO KNOWN AS TYPE
CC CC I LEADER PEPTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AE004511; AAG04157.1; -.

```

```

DR HSSP: P00803; 1B12.
DR MEMOPS: S26.001; -.
DR InterPro: IPR000508; SigPTase.
DR InterPro: IPR000223; SigPTase_S26A.
DR Pfam: PF00461; Peptidase_S26; 1.
DR PRINTS: PR00727; LEADERPTASE.
DR PROSITE: PS00501; SPASE_I.1; 1.
DR PROSITE: PS00760; SPASE_I.2; 1.
DR PROSITE: PS00761; SPASE_I.3; 1.
DR Inner membrane; Transmembrane; Hydrolase; Protease; Complete proteome.
FT DOMAIN 4 22 POTENTIAL.
FT TRANSMEM 59 77 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 78 284 POTENTIAL.
FT ACT_SITE 90 90 PERIPLASMIC (POTENTIAL).
FT ACT_SITE 145 145 BY SIMILARITY.
SQ SEQUENCE 284 AA; 32103 MW; ACA143FA8F90ED2C CRC64;

Query Match 100.0%; Score 24; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. NO. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLVIA 6
|||||
DB 7 LLVIA 12

RESULT 9
LEP_PSEFL
ID LEP_PSEFL. STANDARD; PRT: 284 AA.
AC P26844;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Signal peptidase I (EC 3.4.21.89) (Spase I) (leader peptidase I).
CN LEPB.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 10586;
RX MEDLINE=92189595; PubMed=1546969;
RA Black M.T., Munn J.G.R., Allsop A.E.;
RT "On the catalytic mechanism of prokaryotic leader peptidase I.";
RL Biochem. J. 282:539-543(1992).
CC -I- CATALYTIC ACTIVITY: Cleavage of N-terminal leader sequences from
secreted and periplasmic proteins precursor.
CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S26; ALSO KNOWN AS TYPE
I LEADER PEPTIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X56466; CAA39839.1; -.
CC PIR: S22A14; S22A14.
CC HSSP: P00803; 1B12.
CC MEMOPS: S26.001; -.
DR InterPro: IPR000508; SigPTase.
DR InterPro: IPR000223; SigPTase_S26A.
DR Pfam: PF00461; Peptidase_S26; 1.
DR PRINTS: PR00727; LEADERPTASE.
DR PROSITE: PS00501; SPASE_I.1; 1.
DR PROSITE: PS00760; SPASE_I.2; 1.
DR PROSITE: PS00761; SPASE_I.3; 1.
DR Inner membrane; Transmembrane; Hydrolase; Protease.
KW

```

FT TRANSMEM 4 22 POTENTIAL.  
 FT DOMAIN 23 58 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 59 77 POTENTIAL.  
 FT DOMAIN 78 284 PERIPLASMIC (POTENTIAL).  
 FT ACT\_SITE 90 90 BY SIMILARITY.  
 FT ACT\_SITE 145 145 BY SIMILARITY.  
 SQ SEQUENCE 284 AA; 31903 MW; B1B89D956BDA7891 CRC64;

Query Match 100.0%; Score 24; DB 1; Length 284;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LLLVIA 6  
 |||||  
 Db 7 LLLVIA 12

RESULT 10  
 MNTN\_MYCBO STANDARD; PRT; 415 AA.  
 069443:  
 16-OCT-2001 (rel. 40, Created)  
 16-OCT-2001 (rel. 40, Last sequence update)  
 15-JUN-2002 (rel. 41, Last annotation update)  
 DE Probable manganese transport protein mntH (BRAMP) (Fragment).  
 GN MNTN.  
 OS Mycobacterium bovis.  
 OC Bacteria; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.  
 NCBI\_TaxID=1765;  
 [1]  
 SEQUENCE FROM N.A.  
 RX STRAIN-BG:  
 MEDLINE=98294035; PubMed=9632246;  
 RA Agrinof D.D., Krishna S.;  
 RA "Metal ion homeostasis and intracellular parasitism."  
 RT Mol. Microbiol. 28:403-412(1998).  
 CC -1- FUNCTION: H(+)-STIMULATED, HIGHLY SELECTIVE, MANGANESE UPTAKE  
 SYSTEM (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE NRAMP FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 -----  
 DR EMBL: AJ005699; CA06684.1; -;  
 DR InterPro: IPR001046; Nramp.  
 DR Pfam: PF01566; Nramp; 1.  
 DR PRINTS: PR00447; NATRESASCMP.  
 DR ProDom: PD001861; Nramp; 1.  
 DR TIGRfams: TIGR01197; nramp; 1.  
 KM Transport; Symport; Manganese; Transmembrane.  
 FT NON\_TER 1 1  
 FT TRANSMEM 26 46 POTENTIAL.  
 FT TRANSMEM 53 73 POTENTIAL.  
 FT TRANSMEM 107 127 POTENTIAL.  
 FT TRANSMEM 129 149 POTENTIAL.  
 FT TRANSMEM 164 184 POTENTIAL.  
 FT TRANSMEM 203 223 POTENTIAL.  
 FT TRANSMEM 251 271 POTENTIAL.  
 FT TRANSMEM 292 312 POTENTIAL.  
 FT TRANSMEM 316 336 POTENTIAL.  
 FT TRANSMEM 358 378 POTENTIAL.  
 FT TRANSMEM 399 >415 POTENTIAL.  
 FT NON\_TER 415 415  
 SQ SEQUENCE 415 AA; 43564 MW; A006ECPAFEE3F0DB0 CRC64;  
 Query Match 100.0%; Score 24; DB 1; Length 415;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LLLVIA 6  
 |||||  
 Db 168 LLLVIA 173

RESULT 11  
 MNTN\_MYCCTU STANDARD; PRT; 428 AA.  
 ID MNTN\_MYCCTU  
 AC 005916;  
 DT 15-JUL-1998 (rel. 36, Created)  
 DT 15-JUL-1998 (rel. 36, Last sequence update)  
 DT 15-JUN-2002 (rel. 41, Last annotation update)  
 DE Probable manganese transport protein mntH (BRAMP).  
 GN MNTN OR RV0924C OR MT0951 OR MTCY21C12.18C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.  
 NCBI\_TaxID=1773;  
 [1]  
 SEQUENCE FROM N.A.  
 RX STRAIN-H37RV;  
 MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skellern S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;  
 RA "Deciphering the biology of Mycobacterium tuberculosis from the  
 complete genome sequence."  
 RT Nature 393:537-544(1998).  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 laboratory strains."  
 RT Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: H(+)-STIMULATED, HIGHLY SELECTIVE, MANGANESE UPTAKE  
 SYSTEM (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE NRAMP FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 -----  
 DR EMBL: Z95210; CAB08494.1; -;  
 DR EMBL: AE006981; AKK45198.1; ALT\_INIT.  
 DR TIGR: MT0951; -;  
 DR TubercuList: RV0924C; -;  
 DR InterPro: IPR001046; Nramp.  
 DR Pfam: PF01566; Nramp; 1.  
 DR ProDom: PD001861; Nramp; 1.  
 DR TIGRfams: TIGR01197; nramp; 1.  
 KM Transport; Symport; Manganese; Transmembrane; Complete proteome.  
 FT TRANSMEM 33 53 POTENTIAL.  
 FT TRANSMEM 60 80 POTENTIAL.  
 FT TRANSMEM 114 134 POTENTIAL.  
 FT TRANSMEM 136 156 POTENTIAL.

FT TRANSMEM 171 191 POTENTIAL.  
 FT TRANSMEM 210 230 POTENTIAL.  
 FT TRANSMEM 258 278 POTENTIAL.  
 FT TRANSMEM 299 319 POTENTIAL.  
 FT TRANSMEM 322 343 POTENTIAL.  
 FT TRANSMEM 365 385 POTENTIAL.  
 FT TRANSMEM 406 426 POTENTIAL.  
 SQ SEQUENCE 428 AA; 45004 MW; D90CF98756715CD8 CRC64;

Query Match Best Local Similarity 100.0%; Score 24; DB 1; Length 428;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLLVIA 6  
 DB 175 LLLVIA 180

## RESULT 12

YCA4\_PORPU STANDARD; PRT; 437 AA.  
 P51363:  
 01-OCT-1996 (Rel. 34, Created)  
 01-OCT-1996 (Rel. 34, Last sequence update)  
 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Hypothetical 50.2 kDa protein ycf44 (ORF437).  
 GN YCF44.  
 OS Porphyra purpurea.  
 OC Chloroplast.  
 OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.  
 OX NCBI\_TaxID=2787;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Avonport;  
 RA Reith M.E., Munholland J.;  
 RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast  
 genome.";  
 RL Plant Mol. Biol. Rep. 13:333-335(1995).  
 CC -1- SIMILARITY: BELONGS TO THE YCF44 FAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL: U38804; AAC08249.1;  
 Chloroplast; Hypothetical protein.  
 SQ SEQUENCE 437 AA; 50195 MW; 59F02A977B3A1BD9 CRC64;

Query Match Best Local Similarity 100.0%; Score 24; DB 1; Length 437;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLLVIA 6  
 DB 29 LLLVIA 34

RESULT 13  
 HOFB\_HAEIN STANDARD; PRT; 464 AA.  
 ID HOFB\_HAEIN  
 AC P44622:  
 01-NOV-1995 (Rel. 32, Created)  
 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Protein transport protein hofb homolog.  
 GN HOFB OR HOFB OR H10298.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 " "

OX NCBI\_TaxID=727;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Rd / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7342800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.N., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fritchman J.L., Geoghegan N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 Rd.";  
 RL Science 269:496-512(1995).  
 CC -1- SIMILARITY: BELONGS TO THE PULE/OUTE/EXEE/XPSE/XCPR FAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL: U32715; AAC1962.1;  
 DR TIGR; H10298;  
 DR InterPro: IPR001482; GSPIL.E.  
 DR Pfam: PF00437; GSPIL.E; 1.  
 DR Prodom: PD000739; GSPIL.E; 1.  
 DR PROSITE: PS00662; T2SP\_E; 1.  
 DR Transport; ATP-binding; Complete proteome.  
 FT NP\_BIND 264 271 ATP (POTENTIAL).  
 SQ SEQUENCE 464 AA; 52983 MW; 4923156BE45BD6A2 CRC64;

Query Match Best Local Similarity 100.0%; Score 24; DB 1; Length 464;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LLLVIA 6  
 DB 384 LLLVIA 389

RESULT 14  
 BPL1\_HUMAN STANDARD; PRT; 726 AA.  
 ID BPL1\_HUMAN  
 AC P50747; Q09451;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Biotin-protein ligase (EC 6.3.4.-) (Biotin apo-protein ligase)  
 DE [Includes: Biotin-[methylmalonyl-CoA-carboxyltransferase] ligase  
 DE (EC 6.3.4.9); Biotin-[propionyl-CoA-carboxylase (ATP-hydrolyzing)]  
 DE ligase (EC 6.3.4.10) (Holocarboxylase synthetase) (HCS);  
 DE Biotin-[methylcrotonoyl-CoA-carboxylase] ligase (EC 6.3.4.11);  
 DE Biotin-[acetyl-CoA-carboxylase] ligase (EC 6.3.4.15)].  
 GN HLCS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANT HCS PRO-237.  
 RC TISSUE=Liver;  
 RX MEDLINE=95144167; PubMed=7842009;  
 RA Suzuki Y., Aoki Y., Ishida Y., Chiba Y., Iwamatsu A., Kishino T.,  
 RA Mikawa N., Matsubara Y., Naitawa K.;  
 RT "Isolation and characterization of mutations in the human  
 RT holocarboxylase synthetase cDNA.";  
 RL Nat. Genet. 8:122-128(1994).

[2]  
 RP SEQUENCE FROM N.A.  
 RA Ohira M., Seki N., Nagase T., Suzuki E., Nomura N., Ohara O.,  
 RA Hattori M., Sakaki Y., Eki T., Murekami Y., Saito T., Ichikawa H.,  
 RA Ohki M.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RA Shitaya K., Kudoh J., Minoshima S., Kawasaki K., Nakatch E.,  
 RA Shitani A., Asakawa S., Shimizu N.,  
 RT "Genomic sequencing of 1.2-Mb region on human chromosome 21q22.2";  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [14]  
 RP SEQUENCE FROM N.A., AND VARIANTS HCS P-237; W-508; M-550 AND S-581.  
 RX PubMed=11735028;  
 RA Yang X., Aoki Y., Li X., Sakamoto O., Hiratsuka M., Kure S.,  
 RA Taheri S., Christensen E., Inui K., Kubota M., Ohita M., Ohki M.,  
 RA Kudoh J., Kawasaki K., Shitaya K., Shitani A., Asakawa S.,  
 RA Minoshima S., Shimizu N., Narisawa K., Matsubara Y., Suzuki Y.;  
 RA "Structure of human holocarboxylase synthetase gene and mutation  
 spectrum of holocarboxylase synthetase deficiency.";  
 Hum. Genet. 109:526-534(2001).  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20289799; PubMed=10830953;  
 RA Hattori M., Fujiyama A., Taylor T.D., Matanabe H., Yada T.,  
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Seda E.,  
 RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,  
 RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,  
 RA Reichwald K., Rump A., Schülhabel M., Schudy A., Zimmermann W.,  
 RA Rosenthal A., Kudoh J., Shitaya K., Kawasaki K., Asakawa S.,  
 RA Shitani A., Sakaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,  
 RA Minoshima S., Shimizu N., Nordiek A., Hornischer K., Brandt P.,  
 RA Schärfe M., Schoen O., Desario A., Reichelt J., Kauer G., Blecker H.,  
 RA Ramer J., Beck A., Klages S., Hennig S., Rieselmann B., Dagand E.,  
 RA Weinmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,  
 RA Lehnach H., Reinhardt R., Yaspo M.-L.;  
 RA "The DNA sequence of human chromosome 21.";  
 Nature 405:311-319(2000).  
 RN [16]  
 RP SEQUENCE OF 1-92 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98163741; PubMed=9503011;  
 RA Dahmane N., Alt Ghezala G., Gossset P., Chamoun Z.,  
 RA Dufresne-Zacharia M.C., Lopes C., Rabatel N., Gassanova-Maugenre S.,  
 RA Chetoui Z., Abramowski V., Fayet E., Yaspo M.L., Korn B.,  
 RA Blouin J.L., Lehnach H., Poustka A., Antonarakis S.E., Smet P.M.,  
 RA Creau N., Delabar J.M.;  
 RA "Transcriptional map of the 2.5-Mb CBR-ERG region of chromosome 21  
 involved in down syndrome.";  
 Genomics 48:12-23(1998).  
 RN [17]  
 RP VARIANT HCS PRO-237.  
 RX MEDLINE=96125714; PubMed=8541348;  
 RA Aoki Y., Suzuki Y., Sakamoto O., Li X., Takahashi K., Chitake A.,  
 RA Sakata R., Ohura T., Miyabayashi S., Narisawa K.;  
 RT "Molecular analysis of holocarboxylase synthetase deficiency: a  
 missense mutation and a single base deletion are predominant in  
 Japanese patients.";  
 Biochim. Biophys. Acta 1272:166-174(1995).  
 RN [18]  
 RP VARIANT HCS ASN-571.  
 RX MEDLINE=96414311; PubMed=8817339;  
 RA Dupuis L., Leon-Del-Rio A., Leclerc D., Campeau E., Szeeman L.,  
 RA Saudubray J.M., Herman G., Gibson K.M., Gravel R.A.;  
 RT "Clustering of mutations in the biotin-binding region of  
 holocarboxylase synthetase in biotin-responsive multiple carboxylase  
 deficiency.";  
 Hum. Mol. Genet. 5:1011-1016(1996).  
 RN [19]  
 RP VARIANTS HCS PRO-237 AND MET-550.  
 RX MEDLINE=98057254; PubMed=9396568;  
 RA Aoki Y., Suzuki Y., Li X., Sakamoto O., Chikakawa H., Takita S.,

RA Narisawa K.;  
 RT "Characterization of mutant holocarboxylase synthetase (HCS): a Km for  
 RT biotin was not elevated in a patient with HCS deficiency.";  
 RL Pediatr. Res. 42:849-854(1997).  
 RN [10]  
 RP VARIANT HCS GLU-333; ILE-462; SER-581 AND THR-610 DEL.  
 RX MEDLINE=99204600; PubMed=10190325;  
 RA Aoki Y., Li X., Sakamoto O., Hiratsuka M., Akashi H., Xu L.,  
 RA Briones P., Suomalainen T., Baumgartner E.R., Suzuki Y., Narisawa K.;  
 RT "Identification and characterization of mutations in patients with  
 RT holocarboxylase synthetase deficiency.";  
 RL Hum. Genet. 104:143-148(1999).  
 CC -1- FUNCTION: POSTTRANSLATIONAL MODIFICATION OF SPECIFIC PROTEIN BY  
 CC ATTACHMENT OF BIOTIN. ACTS ON VARIOUS CARBOXYLASES SUCH AS ACETYL-  
 CC COA-CARBOXYLASE, PYRUVATE CARBOXYLASE, PROPYLONL COA CARBOXYLASE,  
 CC AND 3-METHYLCROTONYL COA CARBOXYLASE.  
 CC -1- CATALYTIC ACTIVITY: ATP + biotin + apo-[methylmalonyl-CoA:pyruvate  
 CC carboxyltransferase] = AMP + diphosphate + [methylmalonyl-  
 CC CoA:pyruvate carboxyltransferase].  
 CC -1- CATALYTIC ACTIVITY: ATP + biotin + apo-[propanoyl-CoA:carbon-  
 CC dioxide ligase (ADP-forming)] = AMP + diphosphate + [propanoyl-  
 CC CoA:carbon-dioxide ligase (ADP-forming)].  
 CC -1- CATALYTIC ACTIVITY: ATP + biotin + apo-[3-methylcrotonoyl-  
 CC CoA:carbon-dioxide ligase (ADP-forming)] = AMP + diphosphate + [3-  
 CC methylcrotonoyl-CoA:carbon-dioxide ligase (ADP-forming)].  
 CC -1- CATALYTIC ACTIVITY: ATP + biotin + apo-[acetyl-CoA:carbon-dioxide  
 CC ligase (ADP-forming)] = AMP + diphosphate + [acetyl-CoA:carbon-  
 CC dioxide ligase (ADP-forming)].  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and mitochondrial.  
 CC -1- TISSUE SPECIFICITY: MOSTLY EXPRESSED IN MUSCLE, PLACENTA, IN  
 CC LESSER EXTENT IN THE BRAIN, KIDNEY, PANCREAS, LIVER AND LUNG.  
 CC -1- DISEASE: DEFECTS IN HCS ARE THE CAUSE OF HOCARCOXYLASE  
 CC SYNTHETASE DEFICIENCY (HCS) (OR BIOTIN-RESPONSIVE MULTIPLE  
 CC CARBOXYLASE DEFICIENCY), AN AUTOSOMAL RECESSIVE DISORDER  
 CC CHARACTERIZED BY METABOLIC KETOACIDOSIS, HYPERAMONEMIA, EXCRETION  
 CC OF ABNORMAL ORGANIC ACID METABOLITES, AND DERMATITIS. CLINICAL AND  
 CC BIOCHEMICAL SYMPTOMS IMPROVE DRAMATICALLY WITH ADMINISTRATION OF  
 CC BIOTIN.  
 CC -1- SIMILARITY: WITH E.COLI BIRA AND OTHER EUKARYOTIC BIOTIN--PROTEIN  
 CC LIGASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: D23672; BAA04902.1; -  
 DR EMBL: D67328; BAA13332.1; -  
 DR EMBL: AP000697; BAA89434.1; -  
 DR EMBL: AP000703; BAA89434.1; JOINED.  
 DR EMBL: AP000701; BAA89434.1; JOINED.  
 DR EMBL: AP000698; BAA89434.1; JOINED.  
 DR EMBL: AB063285; BAB68550.1; -  
 DR EMBL: AP001727; BAA95511.1; -  
 DR EMBL: AP001726; BAA95510.1; -  
 DR EMBL: AJ001864; CAA05056.1; -  
 DR Genew: HGNC:4976; HLCS.  
 DR MIM: 253270; -  
 DR InterPro: IPR003142; BPL\_C.  
 DR InterPro: IPR004143; BPL\_LiPA\_LiPB.  
 DR InterPro: IPR004408; Bira\_Ligase.  
 DR Pfam: PF02237; BPL\_C.1.  
 DR Pfam: PF03099; BPL\_LiPA\_LiPB.1.  
 DR TIGRFAMs: TIGR00121; bira\_Ligase.1.  
 KW Ligase; Multifunctional enzyme; Disease mutation.  
 FT SIMILAR 563 609 SOME, TO AVIDIN.  
 FT VARIANT 237 237 L -> P (IN HCS).  
 FT FT /FTID=VAR\_005084.  
 FT VARIANT 333 333 V -> E (IN HCS; <10% ACTIVITY).





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:24:35 ; Search time 20.3182 seconds  
(without alignments)  
60.846 Million cell updates/sec

Title: US-09-543-188A-13  
Perfect score: 24  
Sequence: 1 LLLVIA 6

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.virus:\*  
13: sp.vertibrate:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriap:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	100.0	98	13 Q9DGL8	Q9dgl8 gallus galli
2	24	100.0	107	5 Q9NIO1	Q9n101 trypanosoma
3	24	100.0	110	16 Q8XPU9	Q8xpu9 rai1stonia s
4	24	100.0	128	5 Q9NIO2	Q9n1q2 trypanosoma
5	24	100.0	159	16 Q9PCB9	Q9pcb9 xylella fas
6	24	100.0	173	17 Q8TWT8	Q8twt8 methanopyru
7	24	100.0	183	10 Q9SIE7	Q9sie7 arabidopsis
8	24	100.0	193	8 Q9MIU4	Q9miu4 hyphochytri
9	24	100.0	223	8 Q8WET9	Q8wet9 lyonsia hya
10	24	100.0	224	5 Q44086	Q44086 caenorhabdi
11	24	100.0	232	17 Q9YCB9	Q9ycb9 aeropyrum p
12	24	100.0	239	16 Q9ZBX3	Q9zbx3 streptomyc
13	24	100.0	248	2 Q85820	Q85820 actinobacil
14	24	100.0	281	16 P72696	P72696 synechocyst
15	24	100.0	295	16 Q8ZDP9	Q8zdp9 yersinia pe
16	24	100.0	295	16 Q8UAB6	Q8uab6 agrobacteri

17	24	100.0	296	4 Q9UG33	Q9ug33 homo sapien
18	24	100.0	297	16 P74131	P74131 synechocyst
19	24	100.0	301	16 Q8UG43	Q8ug43 agrobacteri
20	24	100.0	306	2 Q9JSG5	Q9js95 corynebacte
21	24	100.0	311	2 Q9F0G9	Q9f0g9 pseudomonas
22	24	100.0	315	11 Q8VG59	Q8vg59 mus musculu
23	24	100.0	316	5 Q16463	Q16463 caenorhabdi
24	24	100.0	317	11 Q8VFE0	Q8vfe0 mus musculu
25	24	100.0	326	16 Q9ZCD2	Q9zcd2 rickettsia
26	24	100.0	349	10 Q64468	Q64468 arabidopsis
27	24	100.0	361	17 Q9YBP9	Q9ybp9 aeropyrum p
28	24	100.0	368	10 Q9SXN7	Q9sxn7 nicotiana t
29	24	100.0	375	16 Q9CD53	Q9cd53 mycobacteri
30	24	100.0	392	16 Q67832	Q67832 aquifex aeo
31	24	100.0	410	11 Q92320	Q92320 mus musculu
32	24	100.0	424	16 Q9AK38	Q9ak38 streptomyc
33	24	100.0	427	2 Q9EUD4	Q9eud4 pseudomonas
34	24	100.0	442	16 Q8RGE7	Q8rgf7 fusobacteri
35	24	100.0	445	16 Q9A164	Q9a164 streptococc
36	24	100.0	469	2 Q9JRM9	Q9jrm9 actinobacil
37	24	100.0	487	16 Q67873	Q67873 aquifex aeo
38	24	100.0	497	16 Q98M07	Q98m07 rhizobium l
39	24	100.0	509	10 Q40025	Q40025 hordeum vul
40	24	100.0	601	16 Q930C1	Q930c1 rhizobium m
41	24	100.0	654	16 Q9K5Q7	Q9k5q7 bacillus ha
42	24	100.0	672	6 Q9BE74	Q9be74 macaca fasc
43	24	100.0	689	10 Q8S0A9	Q8s0a9 oryza sativ
44	24	100.0	696	2 Q54201	Q54201 streptomyc
45	24	100.0	713	11 Q9EP88	Q9ep88 mus musculu

## ALIGNMENTS

RESULT 1	Q9DGL8	PRELIMINARY;	PRT;	98 AA.
ID	Q9DGL8			
AC	Q9DGL8			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DE	01-MAR-2001 (TREMBLrel. 16, Last annotation update)			
DE	Jun-suppressed chemokine.			
JSC.				
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=EMBROBLAST;			
RA	Hartl M., Bister K.;			
RT	"Suppression of genes in jun-transformed avian fibroblasts."			
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AF285876; AAC00529.1; "			
SQ	SEQUENCE 98 AA; 11564 MW; 30D88E540ADD5B CRC64;			
Query Match	100.0%;	Score 24;	DB 13;	Length 98;
Best Local Similarity	100.0%;	Pred. No. 1.8e+02;		
Matches 6;	Conservative	0;	Mismatches 0;	Indels 0;
QY	1 LLLVIA 6			
DB	9 LLLVIA 14			
RESULT 2	Q9NIO1	PRELIMINARY;	PRT;	107 AA.
ID	Q9NIO1			
AC	Q9NIO1			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			

DE Mucin-like protein.  
GN SMUG.  
OS Trypanosoma cruzi.  
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_TaxId=5693;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CL BRENER;  
RX MEDLINE-20209394; PubMed-10744707;  
RA Di Noia J.M., D'Orso I., Sanchez D.O., Frasch A.C.;  
RT "Aurich elements in the 3'-untranslated region of a new mucin-type  
RT gene family of trypanosoma cruzi confers mRNA instability and  
RT modulates translation efficiency.";  
RL J. Biol. Chem. 275:10218-10227(2000).  
DR EMBL: AF203096; AAF67333.1; -  
SQ SEQUENCE 107 AA; 10811 MW; 7FCFE3307BD58420 CRC64;

Query Match 100.0%; Score 24; DB 5; Length 107;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLLVIA 6  
DB 93 LLLVIA 98

RESULT 3  
Q8XP09 PRELIMINARY; PRT; 110 AA.  
ID Q8XP09;  
AC Q8XP09;  
DT 01-MAR-2002 (TReMBLrel. 20, Created)  
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE Probable transmembrane protein.  
GN RSP1537 OR R504814.  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
OC Ralstonia.  
OX NCBI\_TaxId=305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-GM11000;  
RX MEDLINE-21681879; PubMed-11823852;  
RA Salanoubat M., Ganin S., Artiguenave F., Gouzy J., Mangot L.,  
RA Ariat M., Billaut A., Brottier P., Camus J.C., Catolico L.,  
RA Chandler M., Choins N., Claudel-Renard C., Cunnac S., Demange N.,  
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
RA Sigulier P., Thebault P., Whalen M., Winkler P., Levy M.,  
RA Weissbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";  
RL Nature 415:497-502(2002).  
DR EMBL: AL646085; CAD18688.1; -  
KW Plasmid; Complete proteome.  
SQ SEQUENCE 110 AA; 12295 MW; 3F8A16A7F019E427 CRC64;

Query Match 100.0%; Score 24; DB 16; Length 110;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLLVIA 6  
DB 72 LLLVIA 77

RESULT 4  
Q9NIO2 PRELIMINARY; PRT; 128 AA.  
ID Q9NIO2;  
AC Q9NIO2;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
DE Mucin-like protein.  
SQ SEQUENCE 159 AA; 17586 MW; 4666CA54D543D17 CRC64;

GN SMUG.  
OS Trypanosoma cruzi.  
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
OX NCBI\_TaxId=5693;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CL BRENER;  
RX MEDLINE-20209394; PubMed-10744707;  
RA Di Noia J.M., D'Orso I., Sanchez D.O., Frasch A.C.;  
RT "Aurich elements in the 3'-untranslated region of a new mucin-type  
RT gene family of trypanosoma cruzi confers mRNA instability and  
RT modulates translation efficiency.";  
RL J. Biol. Chem. 275:10218-10227(2000).  
DR EMBL: AF203095; AAF67332.1; -  
SQ SEQUENCE 128 AA; 12746 MW; 2DB356885D8E3715 CRC64;

Query Match 100.0%; Score 24; DB 5; Length 128;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LLLVIA 6  
DB 114 LLLVIA 119

RESULT 5  
Q9PGB9 PRELIMINARY; PRT; 159 AA.  
ID Q9PGB9;  
AC Q9PGB9;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE Hypothetical protein Xf0383.  
GN Xf0383.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xylella.  
OX NCBI\_TaxId=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-9ASC;  
RX MEDLINE-20365717; PubMed-10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvaranga R., Alves L.M.C., Araya J.E., Bal G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Bioness M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
RA Gantier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
RA Klieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montelero-Vicorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,  
RA Pelxoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,  
RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,  
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
RL Nature 406:151-159(2000).  
DR EMBL: AE003890; AAF83193.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 159 AA; 17586 MW; 4666CA54D543D17 CRC64;

Query Match 100.0%; Score 24; DB 16; Length 159;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||  
DB 53 LLLVIA 58

## RESULT 6

Q8TW28 PRELIMINARY; PRT; 173 AA.  
Q8TW28;  
DT 01-JUN-2002 (TREMBLrel. 21, Created)  
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Permease related to cation transporters.  
GN MK0883.  
NCBI\_TaxID=2320;

Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
Methanopyrus

QX NCBI\_TaxID=2320;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AV19 / DSM 6324 / JCM 9639;  
RX MEDLINE=21927647; PubMed=11930014;

RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,  
Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,  
Malykh A.G., Koonin E.V., Kozlovskiy S.A.;  
RA "The complete genome of hyperthermophilic Methanopyrus kandleri AV19  
RT and monophyly of archaeal methanogens.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).

DR EMBL: AE010378; AAM02096.1;  
KW Complete proteome.

SO SEQUENCE 173 AA; 17847 MW; FF7A3981C2A5B2E3 CRC64;

Query Match 100.0%; Score 24; DB 17; Length 173;  
Best Local Similarity 100.0%; Pred. NO. 3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||  
DB 93 LLLVIA 98

## RESULT 7

Q9SIE7 PRELIMINARY; PRT; 183 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Atg22170 protein.  
GN ATG22170.

OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;

RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-CV, COLUMBIA;  
RA MEDLINE=20083487; PubMed=10617197;

RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
Fuji C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
Cronin L.A., Shea M., Vanaken S.E., Umeyam L., Tallon L.J., Gill J.E.,  
Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
Copenhaver G.P., Preuss D., Nielsen W.C., White O., Eisen J.A.,  
RA Salzberg S.L., Fraser C.M., Venter J.C.;  
"Sequence and analysis of chromosome 2 of the plant Arabidopsis

RT thaliana.";  
RL Nature 402:761-768(1999).

RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV, COLUMBIA;  
RA Lin X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC007168; AAD23623.1;  
DR InterPro: IPR001024; Lipoxigenase\_LH2.  
DR Pfam: PF01477; Pfam: 1.  
SO SEQUENCE 183 AA; 20130 MW; B0322292368B92D CRC64;

Query Match 100.0%; Score 24; DB 10; Length 183;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||  
DB 12 LLLVIA 17

## RESULT 8

Q9MI04 PRELIMINARY; PRT; 193 AA.

DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Cytochrome c oxidase subunit II (Fragment).  
GN COX2.

OS Hypochytrium catenoides.  
OC Mitochondrion.

OC Eukaryota; stramenopiles; Hyphochytriomycetes; Hyphochytriales;  
OC Hypochytrium.

OX NCBI\_TaxID=42384;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-ATCC18719;

RA Hudspeeth D.S.S., Nadler S.A., Hudspeeth M.E.S.;  
RT "A cytochrome c oxidase II molecular phylogeny of the  
RL Mycologica 92:674-684(2000)."  
DR EMBL: AF086701; AAF80243.1;  
DR InterPro: IPR001505; Copper\_CuA.  
DR InterPro: IPR002429; Cyt\_c-ox\_2.

DR Pfam: PF00116; COX2; 1.  
DR Pfam: PF02790; COX2\_TM; 1.  
DR PRINTS: PR01166; CYCOXIDASEII.

DR Prodom: PD000131; Copper\_CuA; 1.  
KW Mitochondrion.

FT NON\_TER 1  
FT NON\_TER 193

SO SEQUENCE 193 AA; 22017 MW; D28BD3E4A7A73B67 CRC64;

Query Match 100.0%; Score 24; DB 8; Length 193;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
|||||  
DB 65 LLLVIA 70

## RESULT 9

Q8WET9 PRELIMINARY; PRT; 223 AA.

DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Cytochrome oxidase I (Fragment).  
GN COL.

OS Lyonsia hyalina.

OG Mitochondrion.  
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Anomalodesmata; Pholadomyoida;  
 OC Pandoroidae; Lyonsiidae; Lyonsia.  
 OX NCBI\_TaxID=120442;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Giribet G., Wheeler W.C.;  
 RT "On bivalve phylogeny: a high-level phylogeny of the mollusk class  
 RT Bivalvia based on a combined analysis of morphology and DNA sequence  
 RT data."  
 RL Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF120654; AAL55504.1; -  
 DR InterPro: IPR000883; COX1.  
 DR Pfam: PF00115; COX1; 1.  
 KW Mitochondrion.  
 FT NON\_TER 1  
 FT NON\_TER 223  
 SQ SEQUENCE 223 AA: 24042 MW: 6758BFA6F2D1C725 CRC64;  
 Query Match 100.0%; Score 24; DB 8; Length 223;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 LLLVIA 6  
 DB 179 LLLVIA 184  
 RESULT 10  
 O44086  
 ID O44086; PRELIMINARY; PRT; 224 AA.  
 AC O44086;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE ZK994.4 protein.  
 GN ZK994.4  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodertinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Cosey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,  
 RA Smaildon N., Smith A., Sonhammer E., Steden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Wellstock L., Wilkinson-Sproat J., Wohlman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans."  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Davidson S., Wohlmann P.;  
 RT "The sequence of C. elegans cosmid ZK994."  
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF022977; AAB88611.1; -  
 DR InterPro: IPR001611; LRR.  
 DR InterPro: IPR000483; LRR\_Chern.  
 DR InterPro: IPR000372; LRR\_Nterm.

DR InterPro: IPR003592; LRR\_out.  
 DR Pfam: PF00560; LRR; 3.  
 DR Pfam: PF01463; LRRCT; 1.  
 DR Pfam: PF01462; LRRNT; 1.  
 DR SMART: SM00370; LRR; 3.  
 DR SMART: SM00082; LRRCT; 1.  
 DR SMART: SM00013; LRRNT; 1.  
 SQ SEQUENCE 224 AA: 25885 MW: A634CAE75770BC6A CRC64;  
 Query Match 100.0%; Score 24; DB 5; Length 224;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 LLLVIA 6  
 DB 6 LLLVIA 11  
 RESULT 11  
 O9YCB9  
 ID O9YCB9; PRELIMINARY; PRT; 232 AA.  
 AC O9YCB9;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein APE1337.  
 GN APE1337.  
 OS Aeropyrum pernix.  
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;  
 OC Desulfurococcales; Aeropyrum.  
 OX NCBI\_TaxID=56636;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K1;  
 RX MEDLINE=99310339; PubMed=10382966;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,  
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Akai A., Kosugi H.,  
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
 RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,  
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic hyper-thermophilic  
 RT crenarchaeon, Aeropyrum pernix K1."  
 RL DNA Res. 6:83-101(1999).  
 DR EMBL: AP000061; BAA80329.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 232 AA: 24673 MW: EBECBF8BF27691A0 CRC64;  
 Query Match 100.0%; Score 24; DB 17; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 LLLVIA 6  
 DB 82 LLLVIA 87  
 RESULT 12  
 O9ZBX3  
 ID O9ZBX3; PRELIMINARY; PRT; 239 AA.  
 AC O9ZBX3;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Possible integral membrane protein.  
 GN SCO3959 OR SCD78.26.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M145;

RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajadream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL034355; CAA2231.1; -  
 SQ SEQUENCE 239 AA; 24760 MW; 73165C3F4F575A2 CRC64;

Query Match 100.0%; Score 24; DB 16; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 LLLVIA 6  
 |||||  
 29 LLLVIA 34

RESULT 13  
 085820 PRELIMINARY; PRT; 248 AA.

DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Polc.

GN Polc.  
 OS Actinobacillus actinomycetemcomitans (Haemophilus  
 OS actinomycetemcomitans).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Actinobacillus.  
 ON NCBI\_TaxID=714;  
 RX [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-Y4.  
 RA Winston J.L., Szate S.A., Dyer D.W., McLaughlin R.E.;  
 RT "Demonstration of a polyamine requirement for growth of Actinobacillus  
 RT actinomycetemcomitans, and identification of the putative polyamine  
 RT transport operon."  
 RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.

CC - FUNCTION: PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM.  
 CC - PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE SUBSTRATE ACROSS  
 CC THE MEMBRANE (BY SIMILARITY).  
 CC - SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-  
 CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS.

DR EMBL: AF077856; AAC27497.1; -  
 DR InterPro: IPR000515; BPD\_transp.  
 DR Pfam: PF00528; BPD\_transp. 1.  
 DR PROSITE: PS00402; BPD\_TRANS\_PNN\_MEMBR. 1.  
 KW Transmembrane; Transport.  
 SQ SEQUENCE 248 AA; 27390 MW; 3f641c8ab7356607 CRC64;

Query Match 100.0%; Score 24; DB 2; Length 248;  
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
 |||||  
 DB 232 LLLVIA 237

RESULT 14  
 P72696 PRELIMINARY; PRT; 281 AA.  
 AC P72696;  
 DT 01-FEB-1997 (TREMBlrel. 02, Created)

DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Hypothetical protein slr0241.  
 GN SLR0241.

OS Synecocystis sp. (Strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
 ON NCBI\_TaxID=1148;  
 RX [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE-97061201; PubMed-8905231;

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyaajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M.,  
 RA Tabata S.;

RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synecocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions."  
 RL DNA Res. 3:109-136(1996).

DR EMBL: D90900; BAA16703.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 281 AA; 31695 MW; 08CA1CC1D6D5DA1B CRC64;

Query Match 100.0%; Score 24; DB 16; Length 281;  
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
 |||||  
 DB 264 LLLVIA 269

RESULT 15  
 Q8ZDP9 PRELIMINARY; PRT; 295 AA.

DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative membrane protein.

GN YPO2507.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 ON NCBI\_TaxID=632;  
 RX [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-CO-92 / BIOVAR ORIENTALIS;  
 RX MEDLINE-21470413; PubMed-11586360;

RA Parkhill J., Wren B.W., Thomson N.R., Tittball R.W., Holden M.T.G.,  
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,  
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,  
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 RA Fellwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,  
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 RA Simmonds M., Skellton J., Stevens K., Whitehead S., Barrell B.G.;

RT "Genome sequence of Yersinia pestis, the causative agent of plague."  
 RL Nature 413:523-527(2001).  
 DR EMBL: AJ414152; GAC91312.1; -  
 DR InterPro: IPR000620; DUF6.  
 DR Pfam: PF00892; DUF6. 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 295 AA; 30755 MW; 1174738653C02B7 CRC64;

Query Match 100.0%; Score 24; DB 16; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LLLVIA 6  
 |||||  
 DB 17 LLLVIA 22

Mon Jan 6 15:08:51 2003

us-09-543-188a-13.rspc

Page 6

Search completed: January 3, 2003, 15:31:59  
Job time : 21.3182 secs

---

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:21:45 : Search time 26.5909 Seconds  
(without alignments)  
30.067 Million coil updates/sec

Title: US-09-543-188a-21

Perfect score: 34

Sequence: 1 IQIWF 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Actual number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_101002:\*

1: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*  
23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	100.0	6	23	AAU11844
2	32	94.1	49	22	AAE10210
3	32	94.1	79	21	AAAG06438
4	32	94.1	79	21	AAAG24460
5	32	94.1	79	21	AAAG35187
6	32	94.1	79	21	AAAG35298
7	32	94.1	88	21	AAAG06437
8	32	94.1	88	21	AAAG24459
9	32	94.1	88	21	AAAG35186
10	32	94.1	88	21	AAAG35297

11	32	94.1	101	23	ABG60097
12	32	94.1	110	23	ABG60099
13	32	94.1	126	21	AAAG35185
14	31	91.2	2243	22	ABG29035
15	31	91.2	2478	22	ABG30281
16	30	88.2	51	21	AAAG03171
17	30	88.2	51	22	ABH14713
18	30	88.2	52	23	ABP07377
19	30	88.2	70	20	AA114458
20	30	88.2	74	21	AAI17988
21	30	88.2	98	22	AAU60587
22	30	88.2	122	21	AAI17987
23	30	88.2	123	21	AAI17986
24	30	88.2	305	23	ABP27371
25	30	88.2	305	23	AAU77623
26	30	88.2	603	22	ABG21614
27	30	88.2	739	22	ABG03417
28	30	88.2	739	22	ABG09732
29	30	88.2	739	22	ABG12286
30	29	85.3	24	7	AAPE0356
31	29	85.3	115	22	AAU65123
32	29	85.3	179	22	AAU87351
33	29	85.3	179	23	AAU65354
34	29	85.3	185	22	AAU63958
35	29	85.3	261	22	AAU87428
36	29	85.3	352	22	AAU31812
37	29	85.3	371	22	AAU65431
38	29	85.3	391	19	AAU30680
39	29	85.3	391	19	AAU60255
40	29	85.3	391	19	AAU57324
41	29	85.3	391	20	AAU26166
42	29	85.3	391	22	AAU74887
43	29	85.3	433	20	AAU03770
44	29	85.3	433	21	AAU90622
45	29	85.3	433	21	AAU90656

#### ALIGNMENTS

RESULT 1	
AAU11844	
ID	AAU11844 standard; peptide; 6 AA.
XX	
AC	AAU11844;
DT	
DT	26-MAR-2002 (first entry)
DE	
DE	Peptide ligand for Prion protein, PrP, #19.
KW	Prion protein; PrP; ligand; octapeptide motif; scrapie;
KW	Prion-associated disease; Creutzfeldt-Jakob disease;
KW	Gerstmann-Strausler-Scheinker disease; fatal familial insomnia;
KW	Feline spongiform encephalopathy; bovine spongiform encephalopathy;
KW	transmissible mink encephalopathy; exotic ungulate encephalopathy;
KW	chronic wasting disease.
XX	
OS	Synthetic.
XX	
PN	WO200177687-A2.
PD	18-OCT-2001.
XX	
PF	05-APR-2001; 2001MO-US11150.
XX	
PR	05-APR-2000; 2000US-0543188.
XX	
PA	(VITE-) VI TECHNOLOGIES INC.
XX	
PI	Hammond DJ, Wiltshire VR, Cardonelli R, Shen H;
XX	
DR	WPI; 2002-061944/08.
XX	

Human DTRP polype  
Human DTRP polype  
Zea mays protein f  
Novel human diagno  
Novel human diagno  
Human secreted pro  
Human nervous syst  
Human OREX protein  
Human secreted pro  
Arabidopsis thalia  
Propionibacterium  
Arabidopsis thalia  
Streptococcus poly  
S. agalactiae extr  
Novel human diagno  
Novel human diagno  
Novel human diagno  
Novel human diagno  
Sequence of human  
Propionibacterium  
Human gene 10 enco  
Human albumin fusi  
Human prostate can  
Human gene 10 enco  
Novel human secret  
Drosophila melanog  
Glycerol-3-phospha  
Klebsiella pneumon  
Cytosolic glycerol  
Saccharomyces cere  
Human G-protein co  
Human G protein-co  
Human mutant G pro

PT New ligands for prion proteins, useful for detection or removal or  
PT prions and for treating prion-associated diseases, recognize a specific  
PT octapeptide motif -

PS Claim 16; Page 34; 47pp; English.

CC The invention relates to a ligand of less than 6 kD that binds to a  
CC polypeptide containing the sequence GLYTYRGLYGLINPROHISGLYGLY (A) or an  
CC analogue that is the retro-inverso isomer of (A). The sequence A is  
CC an octapeptide motif from the prion protein (Prp). The ligands are  
CC identified by binding assays with the peptide (A) or peptides containing  
CC (A). The ligands are used for detecting prion proteins (or prions) in  
CC biological or environmental samples, e.g. for diagnosis, also for  
CC removing them from samples, for treating or retarding development of  
CC prion-associated diseases, especially Creutzfeldt-Jakob diseases (in  
CC latrogenic, new variant, familial or sporadic forms), but also  
CC Gerstmann-Strausler-Scheinker disease, fatal familial insomnia,  
CC scrapie, bovine or feline spongiform encephalopathy, transmissible mink  
CC or exotic ungulate encephalopathy, or chronic wasting disease. The  
CC present sequence is a ligand of the invention.

Sequence 6 AA:

Query Match 100.0%; Score 34; DB 23; Length 6;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWIIF 6  
| | | | |  
Db 1 IQIWIIF 6

RESULT 2  
AAE10210

ID AAE10210 standard; peptide; 49 AA.

AC AAE10210;

DT 29-NOV-2001 (first entry)

DE Human bone marrow derived contig peptide, SEQ ID NO: 75.

KW Human; bone marrow; cytokine; cell proliferation; cell differentiation;  
KW stem cell growth; haematopoiesis regulation; immunoregulatory; anaemia;  
KW Parkinson's disease; Alzheimer's disease; amyotrophic lateral sclerosis;  
KW Infection; HIV; malaria; fungal; cancer; autoimmune disorder; arthritis;  
KW Crohn's disease; Inflammatory bowel disease; leukaemia; gene therapy.

XX Homo sapiens.

XX WO200166558-A1.

XX 13-SEP-2001.

XX 07-MAR-2001; 2001WO-US07274.

XX 07-MAR-2000; 2000US-0519705.

XX 19-MAY-2000; 2000US-0574454.

XX 14-JUL-2000; 2000US-0616847.

XX 30-NOV-2000; 2000US-0250583.

XX (HYSE-) HYSEQ INC.

XX Ford JE, Boyle BJ, Tang YT, Liu C, Zhou P, Asundi V, Zhang J;  
PI Drmanac RT;

XX WPI; 2001-565565/63.

XX N-PSDB; AAD17391.

PT Isolated polynucleotide encoding bone marrow derived polypeptides  
PT useful for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis,  
PT Crohn's disease, and inflammatory bowel disease

PS Claim 10; Page 113; 169pp; English.

CC The invention relates to novel human bone marrow polynucleotides and  
CC polypeptides. The polypeptides of the invention may exhibit various  
CC activities e.g. cytokine and cell proliferation/differentiation,  
CC stem cell growth, haematopoiesis regulation and immune stimulation/  
CC suppression. The polynucleotides and polypeptides of the invention may  
CC be used to detect bone marrow cells, for treating, e.g., Parkinson's  
CC disease, Alzheimer's disease, anaemia, amyotrophic lateral sclerosis,  
CC infections such as HIV, malaria and fungal infections, cancer, autoimmune  
CC disorders, arthritis, Crohn's disease, inflammatory bowel disease, and  
CC leukaemia. The polynucleotide may also be used as markers for tissues in  
CC which the corresponding protein is preferentially expressed; as molecular  
CC weight markers and as probes to hybridise and discover novel, related DNA  
CC sequences. The polypeptide may be used in assays to determine biological  
CC activity, to raise antibodies or illicit an immune response and as a  
CC reagent in assays. The present sequence is a human bone marrow derived  
CC contig peptide of the invention.

Sequence 49 AA:

Query Match 94.1%; Score 32; DB 22; Length 49;  
Best Local Similarity 66.7%; Pred. No. 64;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWIIF 6  
: | | | : |  
Db 6 VQIWIIF 11

RESULT 3  
AAG06438

ID AAG06438 standard; Protein; 79 AA.

XX AAG06438;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 3211.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0133247.

XX 04-MAY-1999; 99US-0133248.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 06-MAY-1999; 99US-0132487.



PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0138119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 94.1%; Score 32; DB 21; Length 79;  
Best Local Similarity 83.3%; Pred. No. 1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IQIWF 6  
11111  
18 IQIWL 23

RESULT 4  
AAG24460  
ID AAG24460 standard; Protein; 79 AA.

XX AAG24460;

AC 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 28139.

XX Protein identification: signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

PN 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142554.  
PR 02-JUL-1999; 99US-0142855.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0148171.  
PR 10-AUG-1999; 99US-0148319.  
PR 11-AUG-1999; 99US-0148341.  
PR 12-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 13-AUG-1999; 99US-0149175.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.

PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161992.  
PR 29-OCT-1999; 99US-0162142.  
PR 29-OCT-1999; 99US-0162142.  
  
Query Match 94.1%; Score 32; DB 21; Length 79;  
Best Local Similarity 83.3%; Pred. No. 1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 IQIWF 6  
DB 18 IQIWL 23  
  
RESULT 5  
AAG35187  
ID AAG35187 standard; Protein; 79 AA.  
XX  
AC AAG35187;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Zea mays protein fragment SEQ ID NO: 42948.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.  
XX  
OS Zea mays subsp. mays.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139819.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145813.  
PR 27-JUL-1999; 99US-0145818.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154479.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157573.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 94.1%; Score 32; DB 21; Length 79;  
Best Local Similarity 83.3%; Pred. No. 1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
| | | | |  
Db 18 IQIWL 23

RESULT 6  
AAG35298  
ID AAG35298 standard; Protein; 79 AA.

AC AAG35298;

DT 18-OCT-2000 (first entry)

XX Zea mays protein fragment SEQ ID NO: 43098.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence; corn.

XX Zea mays subsp. mays.

XX EP1033405-A2.

XX PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 18-MAY-1999; 99US-0134370.

PR 19-MAY-1999; 99US-0134768.

PR 20-MAY-1999; 99US-0134941.

PR 21-MAY-1999; 99US-0135124.

PR 24-MAY-1999; 99US-0135353.

PR 25-MAY-1999; 99US-0135629.

PR 27-MAY-1999; 99US-0136021.

PR 28-MAY-1999; 99US-0136592.

PR 01-JUN-1999; 99US-0136782.

PR 03-JUN-1999; 99US-0137222.

PR 04-JUN-1999; 99US-0137528.

PR 07-JUN-1999; 99US-0137502.

PR 08-JUN-1999; 99US-0137724.

PR 10-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144684.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145313.

PR 27-JUL-1999; 99US-0145318.

PR 27-JUL-1999; 99US-0145319.

PR 28-JUL-1999; 99US-0145351.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0146389.

PR 04-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 05-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 06-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148584.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160984.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 94.1%; Score 32; DB 21; Length 79;  
Best Local Similarity 83.3%; Pred. No. 1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IQIWI 6  
Db 18 IQIWI 23

RESULT 7  
AAG06437  
ID AAG06437 standard; Protein; 88 AA.  
AC AAG06437;  
XX 17-OCT-2000 (first entry)  
DT  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 3210.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139482.  
PR 18-JUN-1999; 99US-0139484.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.

```
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155859.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match          94.1%; Score 32; DB 21; Length 88;
Best Local Similarity 93.3%; Pred. No. 1;le=02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 IQIWF 6
Db 27 IQIWL 32

RESULT 8
AAG24459
ID AAG24459 standard; Protein: 88 AA.
XX
AC AAG24459;
```

```
XX 17-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SEQ ID NO: 28138.
DE
XX Protein Identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0125264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.*
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149176.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
```



PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 94.1%; Score 32; DB 21; Length 88;  
Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IOIWF 6  
Db 27 IOIWF 32

RESULT 9  
AAG35186  
ID AAG35186 standard; Protein; 88 AA.

XX AAG35186;

XX 18-OCT-2000 (first entry)

XX zea mays protein fragment SEQ ID NO: 42947.

KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence; corn.  
XX zea mays subsp. mays.  
OS EPI033405-A2.  
XX 06-SEP-2000.  
PD 25-FEB-2000; 2000EP-0301439.  
PF 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147493.  
PR 10-AUG-1999; 99US-0148171.  
PR 10-AUG-1999; 99US-0148319.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.

PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159337.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 94.1%; Score 32; DB 21; Length 88;  
Best Local Similarity 83.3%; Pred. NO. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 IOIWF 6  
|||:|  
Db 27 IOIWF 32

RESULT 10  
AAG35297

ID AAG35297 standard; Protein; 88 AA.

XX AAG35297;

XX 18-OCT-2000 (first entry)

DE Zea mays protein fragment SEQ ID NO: 43097.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.

XX Zea mays subsp. mays.

XX EPI033405-A2.

XX 06-SEP-2000.

PD

XX 25-FEB-2000: 2000EP-0301439.  
XX  
PR 25-FEB-1999: 99US-0121825.  
PR 05-MAR-1999: 99US-0123180.  
PR 09-MAR-1999: 99US-0123548.  
PR 23-MAR-1999: 99US-0125788.  
PR 25-MAR-1999: 99US-0126264.  
PR 29-MAR-1999: 99US-0126785.  
PR 01-APR-1999: 99US-0127462.  
PR 06-APR-1999: 99US-0128234.  
PR 08-APR-1999: 99US-0128714.  
PR 16-APR-1999: 99US-0129845.  
PR 19-APR-1999: 99US-0130077.  
PR 21-APR-1999: 99US-0130449.  
PR 23-APR-1999: 99US-0130510.  
PR 23-APR-1999: 99US-0130891.  
PR 28-APR-1999: 99US-0131449.  
PR 30-APR-1999: 99US-0132048.  
PR 30-APR-1999: 99US-0132407.  
PR 04-MAY-1999: 99US-0132484.  
PR 05-MAY-1999: 99US-0132485.  
PR 06-MAY-1999: 99US-0132486.  
PR 06-MAY-1999: 99US-0132487.  
PR 07-MAY-1999: 99US-0132863.  
PR 11-MAY-1999: 99US-0134218.  
PR 14-MAY-1999: 99US-0134219.  
PR 14-MAY-1999: 99US-0134221.  
PR 14-MAY-1999: 99US-0134570.  
PR 18-MAY-1999: 99US-0134768.  
PR 19-MAY-1999: 99US-0134941.  
PR 20-MAY-1999: 99US-0135124.  
PR 21-MAY-1999: 99US-0135533.  
PR 24-MAY-1999: 99US-0135629.  
PR 25-MAY-1999: 99US-0136021.  
PR 27-MAY-1999: 99US-0136392.  
PR 28-MAY-1999: 99US-0136782.  
PR 01-JUN-1999: 99US-0137222.  
PR 03-JUN-1999: 99US-0137528.  
PR 04-JUN-1999: 99US-0137502.  
PR 07-JUN-1999: 99US-0137724.  
PR 08-JUN-1999: 99US-0138094.  
PR 10-JUN-1999: 99US-0138540.  
PR 10-JUN-1999: 99US-0138847.  
PR 14-JUN-1999: 99US-0139119.  
PR 16-JUN-1999: 99US-0139452.  
PR 16-JUN-1999: 99US-0139453.  
PR 17-JUN-1999: 99US-0139492.  
PR 18-JUN-1999: 99US-0139454.  
PR 18-JUN-1999: 99US-0139455.  
PR 18-JUN-1999: 99US-0139456.  
PR 18-JUN-1999: 99US-0139457.  
PR 18-JUN-1999: 99US-0139458.  
PR 18-JUN-1999: 99US-0139459.  
PR 18-JUN-1999: 99US-0139460.  
PR 18-JUN-1999: 99US-0139461.  
PR 18-JUN-1999: 99US-0139462.  
PR 18-JUN-1999: 99US-0139463.  
PR 18-JUN-1999: 99US-0139750.  
PR 21-JUN-1999: 99US-0139763.  
PR 21-JUN-1999: 99US-0139817.  
PR 22-JUN-1999: 99US-0139817.  
PR 23-JUN-1999: 99US-0140353.  
PR 23-JUN-1999: 99US-0140354.  
PR 24-JUN-1999: 99US-0140695.  
PR 28-JUN-1999: 99US-0140823.  
PR 29-JUN-1999: 99US-0140991.  
PR 30-JUN-1999: 99US-0141287.  
PR 01-JUL-1999: 99US-0141842.  
PR 01-JUL-1999: 99US-0142154.  
PR 02-JUL-1999: 99US-0142055.  
PR 06-JUL-1999: 99US-0142390.  
  
PR 08-JUL-1999: 99US-0142803.  
PR 09-JUL-1999: 99US-0142920.  
PR 12-JUL-1999: 99US-0142977.  
PR 13-JUL-1999: 99US-0143542.  
PR 14-JUL-1999: 99US-0143624.  
PR 15-JUL-1999: 99US-0144005.  
PR 16-JUL-1999: 99US-0144085.  
PR 16-JUL-1999: 99US-0144086.  
PR 19-JUL-1999: 99US-0144325.  
PR 19-JUL-1999: 99US-0144331.  
PR 19-JUL-1999: 99US-0144332.  
PR 19-JUL-1999: 99US-0144333.  
PR 19-JUL-1999: 99US-0144334.  
PR 19-JUL-1999: 99US-0144335.  
PR 20-JUL-1999: 99US-0144352.  
PR 20-JUL-1999: 99US-0144632.  
PR 20-JUL-1999: 99US-0144684.  
PR 21-JUL-1999: 99US-0144814.  
PR 21-JUL-1999: 99US-0145086.  
PR 21-JUL-1999: 99US-0145088.  
PR 22-JUL-1999: 99US-0145085.  
PR 22-JUL-1999: 99US-0145087.  
PR 22-JUL-1999: 99US-0145089.  
PR 22-JUL-1999: 99US-0145192.  
PR 23-JUL-1999: 99US-0145145.  
PR 23-JUL-1999: 99US-0145218.  
PR 23-JUL-1999: 99US-0145224.  
PR 26-JUL-1999: 99US-0145276.  
PR 27-JUL-1999: 99US-0145913.  
PR 27-JUL-1999: 99US-0145918.  
PR 27-JUL-1999: 99US-0145919.  
PR 28-JUL-1999: 99US-0145951.  
PR 02-AUG-1999: 99US-0146386.  
PR 02-AUG-1999: 99US-0146388.  
PR 02-AUG-1999: 99US-0146389.  
PR 03-AUG-1999: 99US-0147038.  
PR 04-AUG-1999: 99US-0147204.  
PR 04-AUG-1999: 99US-0147302.  
PR 05-AUG-1999: 99US-0147192.  
PR 05-AUG-1999: 99US-0147260.  
PR 06-AUG-1999: 99US-0147303.  
PR 06-AUG-1999: 99US-0147416.  
PR 09-AUG-1999: 99US-0147493.  
PR 09-AUG-1999: 99US-0147935.  
PR 10-AUG-1999: 99US-0148171.  
PR 11-AUG-1999: 99US-0148319.  
PR 12-AUG-1999: 99US-0148341.  
PR 13-AUG-1999: 99US-0148565.  
PR 13-AUG-1999: 99US-0148684.  
PR 16-AUG-1999: 99US-0149368.  
PR 17-AUG-1999: 99US-0149175.  
PR 18-AUG-1999: 99US-0149426.  
PR 20-AUG-1999: 99US-0149722.  
PR 20-AUG-1999: 99US-0149723.  
PR 20-AUG-1999: 99US-0149929.  
PR 23-AUG-1999: 99US-0149902.  
PR 23-AUG-1999: 99US-0149930.  
PR 23-AUG-1999: 99US-0150566.  
PR 26-AUG-1999: 99US-0150884.  
PR 27-AUG-1999: 99US-0151065.  
PR 27-AUG-1999: 99US-0151066.  
PR 27-AUG-1999: 99US-0151080.  
PR 30-AUG-1999: 99US-0151303.  
PR 31-AUG-1999: 99US-0151438.  
PR 01-SEP-1999: 99US-0151930.  
PR 07-SEP-1999: 99US-0152363.  
PR 10-SEP-1999: 99US-0153070.  
PR 13-SEP-1999: 99US-0153758.  
PR 15-SEP-1999: 99US-0154018.  
PR 16-SEP-1999: 99US-0154039.  
PR 20-SEP-1999: 99US-0154779.  
PR 22-SEP-1999: 99US-0155139.  
PR 23-SEP-1999: 99US-0155486.

PR	28-SEP-1999;	99US-015659.
PR	28-SEP-1999;	99US-015645B.
PR	29-SEP-1999;	99US-015659B.
PR	04-OCT-1999;	99US-015711.
PR	05-OCT-1999;	99US-015715C.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0156029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

DT 30-JUL-2002 (first entry)  
XX  
DE Human DITHP polypeptide #157.  
XX  
XX Human; DITHP; diagnostic and therapeutic polypeptide; bone; testis; skin;  
KW cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;  
KW inflammatory disorder; viral infection; bacterial infection; seizure;  
KW fungal infection; parasitic infections; developmental disorder; breast;  
KW endocrine disorder; metabolic disorder; neurological disorder; cervix;  
KW gastrointestinal disorder; transport disorder; gene therapy; kidney;  
KW adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;  
KW thymus.  
XX  
OS Homo sapiens.  
XX  
PN WO200220754-A2.  
XX  
PD 14-MAR-2002.  
XX  
XX 29-AUG-2001; 2001WO-US27127.  
PR 05-SEP-2000; 2000US-229747P.  
PR 05-SEP-2000; 2000US-228748P.  
PR 05-SEP-2000; 2000US-229749P.  
PR 05-SEP-2000; 2000US-229750P.  
PR 05-SEP-2000; 2000US-229751P.  
PR 05-SEP-2000; 2000US-230583P.  
PR 06-SEP-2000; 2000US-230505P.  
PR 06-SEP-2000; 2000US-230514P.  
PR 06-SEP-2000; 2000US-230515P.  
PR 06-SEP-2000; 2000US-230517P.  
PR 06-SEP-2000; 2000US-230518P.  
PR 06-SEP-2000; 2000US-230519P.  
PR 06-SEP-2000; 2000US-230595P.  
PR 06-SEP-2000; 2000US-230597P.  
PR 06-SEP-2000; 2000US-230598P.  
PR 06-SEP-2000; 2000US-230599P.  
PR 06-SEP-2000; 2000US-230610P.  
PR 06-SEP-2000; 2000US-230655P.  
PR 06-SEP-2000; 2000US-230988P.  
PR 07-SEP-2000; 2000US-230951P.  
PR 07-SEP-2000; 2000US-231163P.  
PR 07-SEP-2000; 2000US-231167P.  
XX  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
PI Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup NS, Hillman JL;  
Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;  
Moriyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;  
Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;  
Marwaha R, Chen AJ, Chang SC, Au AP, Imman RR;  
XX  
XX WPI: 2002-383054/41.  
DR N-PSDB; ABR71690.  
XX  
XX  
PT An isolated polynucleotide useful in diagnostics and therapeutics -  
XX  
XX Claim 29: Page 622; 686pp; English.  
XX  
XX The invention relates to human diagnostic and therapeutic (dithp)  
CC polynucleotides and their associated polypeptides (DITHP polypeptides).  
CC The sequences of the invention are used in the treatment and diagnosis of  
CC cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers  
CC (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,  
CC cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or  
CC thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,  
CC psoriasis, osteoporosis), viral infections, bacterial infections, fungal  
CC infections, parasitic infections, developmental disorders (e.g. anaemia,  
CC epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),  
CC endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders  
CC (e.g. obesity, diabetes), neurological disorders (e.g. stroke,  
CC amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal  
CC disorders (e.g. ulcerative colitis, lysinuria) and transport disorders

CC (e.g. myotonic dystrophy, catatonias, peripheral neuropathy). Sequences  
CC ABC59943-ABC60220 represent human DITHP polypeptides of the invention.  
XX  
SQ Sequence 110 AA:  
Query Match 94.1%; Score 32; DB 23; Length 110;  
Best Local Similarity 83.3%; Pred. No. 1.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 1 IQIWIIF 6  
DB 57 IQIWIIF 62  
RESULT 13  
AAG35185  
ID AAG35185 standard; Protein; 126 AA.  
XX  
AC AAG35185;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Zea mays' protein fragment SEQ ID NO: 42946.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.  
XX  
OS Zea mays subsp. mays.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0123788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0131448.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134376.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145216.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148177.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 13-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 94.1%; Score 32; DB 21; Length 126;  
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IQIWMF 6  
||||:|  
Db 65 IQIWMF 70

RESULT 14  
ABG29035  
ID ABG29035 standard; Protein: 2243 AA.

AC ABG29035;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #29026.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN MO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; AAS93222.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity

PS Claim 20; SEQ ID No 59394; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX imaging of sites expressing (II). (I) and (II) are useful for treating  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABG30377 represent novel human  
XX diagnostic amino acid sequences of the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 2243 AA;

Query Match 91.2%; Score 31; DB 22; Length 2243;  
Best Local Similarity 83.3%; Pred. No. 4.2e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWMF 6  
||||:|  
Db 328 IQIWMF 333

RESULT 15  
ABG30281  
ID ABG30281 standard; Protein: 2478 AA.

AC ABG30281;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #30272.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN MO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR N-PSDB; AAS94468.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity

PS Claim 20; SEQ ID No 60640; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX imaging of sites expressing (II). (I) and (II) are useful for treating  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABG30377 represent novel human  
XX diagnostic amino acid sequences of the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 2478 AA;

Query Match 91.2%; Score 31; DB 22; Length 2478;  
Best Local Similarity 83.3%; Pred. No. 4.6e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IQIWMF 6  
||||:|

Mon Jan 6 15:08:52 2003

us-09-543-188a-21.rag

Page 18

Db 563 IQIWF 568

Search completed: January 3, 2003, 15:28:40  
Job time : 28.5909 secs

---



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:00 : Search time 9.95455 Seconds  
(without alignments)  
57.944 Million cell updates/sec

Title: US-09-543-188a-21

Perfect score: 34

Sequence: 1 IQIWIIF 6

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

al number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	97.1	282	2 C58933	ribosomal protein
2	33	97.1	285	2 T24123	hypothetical prote
3	33	97.1	360	2 G83831	spore germination
4	32	94.1	86	2 F85354	small nuclear ribo
5	32	94.1	88	2 A84568	probable small nuc
6	32	94.1	495	2 T12401	NADH2 dehydrogenas
7	31	91.2	394	2 E81286	probable polysacch
8	30	88.2	218	1 R3NT3	ribosomal protein
9	30	88.2	218	1 R3SP3	ribosomal protein
10	30	88.2	220	2 S78395	ribosomal protein
11	30	88.2	229	2 E70340	glutaredoxin-like
12	30	88.2	240	2 S59085	ribosomal protein
13	30	88.2	309	2 T34080	hypothetical prote
14	30	88.2	347	2 T50963	related to cell pr
15	30	88.2	488	2 C59240	type II site-speci
16	29	85.3	144	2 G84935	hypothetical prote
17	29	85.3	217	1 R3LV3	ribosomal protein
18	29	85.3	222	2 B95207	aquaporin [impor
19	29	85.3	222	2 B98072	aquaporin Z, water
20	29	85.3	224	2 S38590	ribosomal protein
21	29	85.3	231	2 T07358	ribosomal protein
22	29	85.3	239	1 R3R23	ribosomal protein
23	29	85.3	240	2 S77496	ribosomal protein
24	29	85.3	243	2 E97593	30S ribosomal prot
25	29	85.3	243	2 A82815	30S ribosomal prot
26	29	85.3	244	2 T11241	Sect-independent t
27	29	85.3	252	2 B41839	ribosomal protein
28	29	85.3	327	2 S76951	hypothetical prote
29	29	85.3	348	2 D70195	hypothetical prote

30	29	85.3	373	2 P00042	actavin - fruit fl
31	29	85.3	391	2 S40059	glycerol-3-phospha
32	29	85.3	583	2 T02382	hypothetical prote
33	29	85.3	720	2 T02361	hypothetical prote
34	29	85.3	920	1 P4NCP	H+-exporting ATPas
35	29	85.3	978	1 RCBY13	regulatory protein
36	28	82.4	71	2 T51537	hypothetical prote
37	28	82.4	90	2 T27041	hypothetical prote
38	28	82.4	92	2 A32127	small nuclear ribo
39	28	82.4	122	2 A31090	hypothetical 14k p
40	28	82.4	122	2 AC1808	ATP synthase chain
41	28	82.4	205	2 PC2191	ribosomal protein
42	28	82.4	219	1 R3KT3	ribosomal protein
43	28	82.4	227	2 T11185	H+-transporting tw
44	28	82.4	227	2 F90622	ATP synthase F0 ch
45	28	82.4	228	2 G83115	30S ribosomal prot

#### ALIGNMENTS

```
RESULT 1
C58933
ribosomal protein S3 - Cyanidioschyzon merolae mitochondrion
C:Species: mitochondrion Cyanidioschyzon merolae
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #extl_change 05-Nov-1999
C:Accession: C58933
R:Ohta, N.; Sato, N.; Kuroiwa, T.
Nucleic Acids Res. 26, 5190-5198, 1998
A:Title: Structure and organization of the mitochondrial genome of the unicellular re
A:Reference number: A58930; MUID:99030526; PMID:9801318
A:Accession: C58933
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-282 <ARN>
A:Cross-references: GB:D89861; NID:g4115781; PIDN:BAA36537.1; PID:d1037523; PID:g4115
C:Genetics:
A:Gene: rps3
C:Keywords: mitochondrion

Query Match
Best local Similarity 97.1%; Score 33; DB 2; Length 282;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWIIF 6
Db 275 IQIWIIF 280

RESULT 2
T24123
hypothetical protein R10D12.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #extl_change 17-Mar-2000
C:Accession: T24123
R:Percy, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19842
A:Accession: T24123
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-285 <WII>
A:Cross-references: EMBL:Z81109; PIDN:CAB03254.1; GSPDB:GN00023; CESP:R10D12.11
A:Experimental source: clone R10D12
C:Genetics:
A:Gene: CESP:R10D12.11
A:Map position: 5
A:Introns: 51/2; 165/3; 240/3
C:Superfamily: Caenorhabditis hypothetical protein C4967.2

Query Match
Best local Similarity 97.1%; Score 33; DB 2; Length 285;
```

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IQIWF 6  
:|||||  
Db 48 VOIWF 53

RESULT 3  
G83831  
spore germination protein BH1455 [imported] - *Bacillus halodurans* (strain C-125)  
C:Species: *Bacillus halodurans*  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: G83831

R.Takemi, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai  
Nucleic Acids Res. 28, 4317-4331, 2000  
A>Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: G83831

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-360 <STO>  
A:Cross-references: GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BAB05174.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH1455

Query Match 97.1%; Score 32; DB 2; Length 360;  
Best Local Similarity 83.3%; Pred. No. 59;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
:|||||  
Db 109 IQIWF 114

RESULT 4  
F85354

small nuclear ribonucleoprotein homolog [imported] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: F85354  
Ranomymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
Nature 402, 769-777, 1999  
A>Title: Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.  
A:Reference number: A85001; MUID:20083488; PMID:10617198  
A:Accession: F85354

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-86 <STO>  
A:Cross-references: GB:NC\_001268; NID:g7269933; PIDN:CAB81026.1; GSPDB:GN00140  
C:Genetics:  
A:Map position: 4  
A:Gene: AT4G30330

Query Match 94.1%; Score 32; DB 2; Length 86;  
Best Local Similarity 83.3%; Pred. No. 23;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
:|||||  
Db 27 IQIWF 32

RESULT 5  
A84568

probable small nuclear ribonucleoprotein E [imported] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: A84568  
R.Llin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Mofeit, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-766, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: A84568  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-88 <STO>  
A:Cross-references: GB:AE002093; NID:g4185140; PIDN:AAD08943.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g18740  
A:Map position: 2

Query Match 94.1%; Score 32; DB 2; Length 88;  
Best Local Similarity 83.3%; Pred. No. 23;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
:|||||  
Db 27 IQIWF 32

RESULT 6  
T12401

MADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - *Sarcophyton glaucum* mitochondrion  
C:Species: *Sarcophyton glaucum*  
C:Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 03-Jun-2002  
C:Accession: T12401  
R.Beaton, M.J.; Roger, A.J.; Cavaller-Smith, T.  
J. Mol. Evol. 47, 697-708, 1998

A>Title: Sequence analysis of the mitochondrial genome of *Sarcophyton glaucum*: Conser  
A:Reference number: Z17505; MUID:99065763; PMID:9847412  
A:Accession: T12401  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-495 <BEA>

A:Cross-references: EMBL:AF064823; NID:g4091912; PID:g4091917; PIDN:AAC99651.1  
C:Genetics:  
A:Genome: mitochondrion  
A:Note: ND4  
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4  
C:Keywords: membrane-associated complex; mitochondrion; NMD; oxidative phosphorylation

Query Match 94.1%; Score 32; DB 2; Length 495;  
Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
:|||||  
Db 219 IQIWF 224

RESULT 7  
E81286

probable polysaccharide modification protein Cj1413c [imported] - *Campylobacter jejuni*  
C:Species: *Campylobacter jejuni*  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
C:Accession: E81286  
P.Parthill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chli  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; Vanvliet, A.; Whitehead, S.; Bar  
Nature 403, 665-668, 2000  
A>Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals  
A:Reference number: A81250; MUID:20150912; PMID:10688204  
A:Accession: E81286

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <PAR>  
A:Cross-references: GB:AL139078; GB:AL111168; NID:g6968723; PIDN:CAB73837.1; PID:g696  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj1413c  
C:Superfamily: kpsd protein

Query Match 91.2%; Score 31; DB 2; Length 394;  
Best Local Similarity 83.3%; Pred. No. 1.6e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IQIWF 6  
|:||||  
Db 105 IEIWF 110

RESULT 8  
R3NT3

ribosomal protein S3, chloroplast - common tobacco chloroplast  
C:Species: chloroplast Nicotiana tabacum (common tobacco)

C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 30-Jan-1998  
C:Accession: A02702; E25943

R: Sugita, M.

submitted to the EMBL Data Library, August 1986  
A:Reference number: A00149

A:Accession: A02702

A:Molecule type: DNA

A:Residues: 1-218 <SUG>  
Experimental source: cv. Bright Yellow 4

Shinozaki, K.; Ohme, M.; Tanaka, M.; Makasugi, T.; Hayashida, N.; Matsubayashi, T.; Zä  
eno, H.; Kamogashira, T.; Yamada, K.; Kusuda, J.; Takaiwa, F.; Kato, A.; Tondoh, N.; Sh  
EMBO J. 5, 2043-2049, 1986

A:Title: The complete nucleotide sequence of the tobacco chloroplast genome: its gene on  
A:Reference number: A38013

A:Contents: annotation; gene organization, sites, features  
R: Tanaka, M.; Makasugi, T.; Sugita, M.; Shinozaki, K.; Sugita, M.

Proc. Natl. Acad. Sci. U.S.A. 83, 6030-6034, 1986

A:Title: Genes for the eight ribosomal proteins are clustered on the chloroplast genome  
A:Reference number: A94118; M01D:86287388; PMID:3016736

A:Accession: E25943

A:Molecule type: DNA

A:Residues: 1-218 <TRAN>

C:Genetics:

A:Gene: rps3

A:Genome: chloroplast

C:Superfamily: Escherichia coli ribosomal protein S3

C:Keywords: chloroplast; protein biosynthesis; ribosome

Query Match  
Best Local Similarity 88.2%; Score 30; DB 1; Length 218;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
|:||||  
Db 209 IKIWF 214

ULF 9  
P3

ribosomal protein S3, chloroplast - spinach chloroplast  
C:Species: chloroplast Spinacia oleracea (spinach)

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999  
C:Accession: S01978

R: Zhou, D.X.; Quigley, F.; Massenot, O.; Mache, R.  
Mol. Gen. Genet. 216, 439-445, 1989

A:Title: Cotranscription of the S10- and spc-like operons in spinach chloroplasts and id  
A:Reference number: S01976; M01D:89313684; PMID:2747623

A:Accession: S01978

A:Molecule type: DNA

A:Residues: 1-218 <ZHO>

A:Cross-references: EMBL:X13336; NID:g12307; PIDN:CAA31715.1; PID:g12310

C:Genetics:

A:Gene: rps3

A:Genome: chloroplast

C:Superfamily: Escherichia coli ribosomal protein S3

C:Keywords: chloroplast; protein biosynthesis; ribosome

Query Match  
Best Local Similarity 88.2%; Score 30; DB 1; Length 218;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6

Db 209 IKIWF 214  
|:||||

RESULT 10  
S78395

ribosomal protein S3, plastid - beechdrops plastid  
C:Species: plastid Epifagus virginiana (beechdrops)

C:Date: 25-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 13-Aug-1999  
C:Accession: S78395

R: Wolfe, K.H.; Morden, C.W.; Ems, S.C.; Palmer, J.D.  
J. Mol. Evol. 35, 304-317, 1992

A:Title: Rapid evolution of the plastid translational apparatus in a nonphotosynthetic  
A:Reference number: S78378; M01D:93021155; PMID:1404416

A:Accession: S78395

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-220 <MOL>  
A:Cross-references: EMBL:M81884; NID:g336917; PIDN:AA65864.1; PID:g336936

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992  
C:Genetics:

A:Gene: rps3

A:Genome: plastid

C:Superfamily: Escherichia coli ribosomal protein S3

C:Keywords: plastid; protein biosynthesis; ribosome

Query Match  
Best Local Similarity 88.2%; Score 30; DB 2; Length 220;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
|:||||  
Db 211 IKIWF 216

RESULT 11  
E70340

glutaredoxin-like protein - Aquifex aeolicus  
C:Species: Aquifex aeolicus

C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
C:Accession: E70340

R: Beckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;  
V.

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70500; M01D:98196666; PMID:9537320

A:Accession: E70340

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-229 <AOE>

A:Cross-references: GB:AE000690; M01D:g2983100; PIDN:AMC06709.1; PID:g2983104; GB:AE00  
A:Experimental source: strain VF5

C:Genetics:

A:Gene: gua

Query Match  
Best Local Similarity 88.2%; Score 30; DB 2; Length 229;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
|:||||  
Db 138 IEIWF 143

RESULT 12  
S59085

ribosomal protein S3, mitochondrial - red alga (Chondrus crispus) mitochondrion  
C:Species: mitochondrial Chondrus crispus (carrageen)

C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 21-Jul-2000  
C:Accession: S59085

R: Leblanc, C.; Boyen, C.; Richard, O.; Bonnard, G.; Grienenberger, J.M.; Kloareg, B.  
J. Mol. Biol. 250, 484-495, 1995

A:Title: Complete sequence of the mitochondrial DNA of the rhodophyte Chondrus crispus

A:Reference number: S59078; MUID:95341681; PMID:7616569  
 A:Accession: S59085  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-240 <LEB>  
 A:Cross-references: EMBL:Z47547; NID:g1019057; PIDN:CA87601.1; PID:g1334481  
 A:Experimental source: female gametophytes  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995  
 C:Genetics:  
 A:Gene: rps3  
 A:Genome: mitochondrion  
 A:Genetic code: SGC3  
 C:Function:  
 A:Pathway: protein biosynthesis  
 C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match 88.2%; Score 30; DB 2; Length 240;  
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

1 QIWIIF 6  
 :|:|:  
 Db 220 VQVWLF 225

RESULT 13  
 T34080  
 hypothetical protein C02F12.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T34080  
 R:Miller, N.  
 submitted to the EMBL Data Library, November 1995  
 A:Description: The sequence of C. elegans cosmid C02F12.  
 A:Reference number: Z21473  
 A:Accession: T34080  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-309 <MIL>  
 A:Cross-references: EMBL:U41545; PIDN:AAA83189.1; CESP:C02F12.1  
 C:Genetics:  
 A:Gene: CESP:C02F12.1  
 A:introns: 23/3; 58/3; 83/3; 140/3; 178/3; 257/3; 285/3

Query Match 88.2%; Score 30; DB 2; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2 QIWIIF 6  
 :|:|:  
 Db 270 QIWIIF 274

RESULT 14  
 T50963  
 related to cell protein precursor [imported] - Neurospora crassa  
 N:Alternate names: protein B24P7.180  
 C:Species: Neurospora crassa  
 C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
 C:Accession: T50963  
 R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
 submitted to the Protein Sequence Database, July 2000  
 A:Reference number: Z25286  
 A:Accession: T50963  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-347 <SCH>  
 A:Cross-references: EMBL:AL389890; GSPDB:GN00116; NCSP:B24P7.180  
 A:Experimental source: BAC clone B24P7; strain OR74A  
 C:Genetics:  
 A:Gene: NCSP:B24P7.180  
 A:Map position: 6

Query Match 88.2%; Score 30; DB 2; Length 347;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+02;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIWIIF 6  
 :|:|:  
 Db 206 VQVWLF 211

RESULT 15  
 C59240  
 type II site-specific deoxyribonuclease (EC 3.1.21.4) HgaI - Haemophilus paragaillinar  
 C:Species: Haemophilus paragaillinarum  
 C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 20-Jun-2000  
 C:Accession: C59240  
 R:Sugisaki, H.  
 Bull. Inst. Chem. Res. Kyoto Univ. 71, 338-342, 1993  
 A:Title: Nucleotide sequence of the gene of HgaI restriction endonuclease.  
 A:Reference number: A59239  
 A:Accession: C59240  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-488 <SUG>  
 A:Cross-references: GB:D17388; NID:g393310; PIDN:BA04208.1; PID:g393313  
 A:Experimental source: strain NCTC 3438  
 C:Superfamily: Haemophilus paragaillinarum type II site-specific deoxyribonuclease Hga  
 C:Keywords: hydrolase

Query Match 88.2%; Score 30; DB 2; Length 488;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QIWIIF 6  
 :|:|:  
 Db 11 QIWIIF 15

Search completed: January 3, 2003, 15:33:18  
 Job time : 10.9545 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:22:10 : Search time 5.18182 Seconds  
(without alignments)  
48.025 Million cell updates/sec

Title: US-09-543-188A-21

Sequence: 1 IQIWI 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: SWISSPROT\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	94.1	224	1	MTGA_ACTICA
2	30	88.2	218	1	RR3_ARATH
3	30	88.2	218	1	RR3_SPTOL
4	30	88.2	218	1	RR3_TOBAC
5	30	88.2	220	1	RR3_ERIYI
6	30	88.2	240	1	RR3_CHOCR
7	30	88.2	277	1	YL11_CAEBL
8	30	88.2	488	1	T2G1_HAEGA
9	29	85.3	144	1	Y052_BUCAI
10	29	85.3	213	1	RR3_MESVI
11	29	85.3	217	1	RR3_MARPO
12	29	85.3	224	1	RR3_MAIZE
13	29	85.3	231	1	RR3_CHLVU
14	29	85.3	239	1	RR3_ORYSA
15	29	85.3	239	1	RR3_SYNY3
16	29	85.3	252	1	RR3_ASTYP
17	29	85.3	252	1	RR3_PHY52
18	29	85.3	324	1	CP21_HORVU
19	29	85.3	391	1	GPDI_YEAST
20	29	85.3	433	1	GP22_HUMAN
21	29	85.3	920	1	PMAL_NEUCR
22	29	85.3	978	1	STR3_YEAST
23	28	82.4	92	1	RUXE_HUMAN
24	28	82.4	122	1	ATPZ_ANASP
25	28	82.4	205	1	RR3_BUCAR
26	28	82.4	216	1	RR3_GUITH
27	28	82.4	219	1	RR3_CYAPA
28	28	82.4	232	1	RR3_ECOLI
29	28	82.4	233	1	RR3_BUCAI
30	28	82.4	234	1	RR3_ACTAC
31	28	82.4	234	1	RR3_HAEIN
32	28	82.4	234	1	RR3_HELPJ
33	28	82.4	234	1	RR3_HELPY

34	28	82.4	243	1	RS3_SYNP6	O24695 synechococ
35	28	82.4	248	1	RS3_ACHSP	P41118 acholoplasm
36	28	82.4	251	1	RS3_PHY51	O66095 phytoplasm
37	28	82.4	270	1	VSP1_ARATH	O49195 arabidopsis
38	28	82.4	313	1	FIXB_ECOLI	P31574 escherichia
39	28	82.4	326	1	YE09_SYNY3	P73594 synechocyst
40	28	82.4	340	1	CLN1_PIG	O95002 sus scrofa
41	28	82.4	374	1	YFIB_SPTCI	P27712 spiropilasma
42	28	82.4	509	1	CCBS_MARPO	P36180 marchantia
43	28	82.4	515	1	NCAP_P134H	P06159 human para
44	28	82.4	551	1	NCAP_P134H	P17241 human para
45	28	82.4	712	1	RR3_CHLRE	O08365 chlamydomon

## ALIGNMENTS

## RESULT 1

ID	MTGA_ACTICA	STANDARD:	PRT:	224 AA.
AC	O24649:			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Monofunctional biosynthetic peptidoglycan transglycosylase (EC 2.4.2.-) (Monofunctional Itase).			
GN	MTGA.			
OS	Acinetobacter calcoaceticus.			
OC	Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;			
OC	Acinetobacter.			
OX	NCBI_TaxID=471;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BD413 / ADP1;			
RX	MEDLINE=95400495; PubMed=7670642;			
RA	Geissdoerfer W., Frosch C.S., Haspel G., Eht S., Hillen W.;			
RT	"Two genes encoding proteins with similarities to rubredoxin and rubredoxin reductase are required for conversion of dodecane to lauric acid in Acinetobacter calcoaceticus ADP1.";			
RL	Microbiology 141:1425-1432(1995).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BD413 / ADP1;			
RX	MEDLINE=97228433; PubMed=9074511;			
RA	Geissdoerfer W., Ratajczak A., Hillen W.;			
RT	"Nucleotide sequence of a putative periplasmic Mn superoxide dismutase from Acinetobacter calcoaceticus ADP1.";			
RL	Gene 186:305-308(1997).			
CC	- FUNCTION: CELL WALL FORMATION (BY SIMILARITY).			
CC	- PATHWAY: FINAL STAGES IN PEPTIDOGLYCAN SYNTHESIS (BY SIMILARITY).			
CC	- SUBCELLULAR LOCATION: Membrane-associated (potential).			
CC	- SIMILARITY: TO THE N-TERMINAL OF BACTERIAL CLASS IA PENICILLIN-BINDING PROTEINS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: 246863; CAA86932.1; -			
DR	InterPro: IPR001264; GT_51.			
DR	Pfam: PF00912; Transglycosyl; 1.			
DR	Prodom: PD001895; GT_51; 1.			
KW	peptidoglycan synthesis; Cell wall; Transferase; Transmembrane.			
FT	TRANSMEM 8 30 POTENTIAL.			
SO	SEQUENCE 224 AA; 26712 MW; F1ESF114FC14E2FE CRC64;			
Query Match	94.1%; Score 32; DB 1; Length 224;			
Best Local Similarity	83.3%; Pred. No. 26;			
Matches	5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			

QY 1 IQIWI 6  
1:||||  
DB 19 IQIWI 24

## RESULT 2

RR3\_ARATH STANDARD: PRT: 218 AA.  
ID P56798;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE Chloroplast 30S ribosomal protein S3.  
GN RPS3.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
NCBI\_TaxID=3702;  
[1]

## SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;  
RX MEDLINE=20039611; PubMed=10574454;  
RA Sato S., Nakamura Y., Kaneko T., Asamizu E., Tabata S.;  
RT "Complete structure of the chloroplast genome of Arabidopsis  
thaliana";  
RL DNA Res. 6:283-290(1999).  
CC -1- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.  
CC -1- SIMILARITY: CONTAINS 1 KH TYPE-2 DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL: AP000423; BAA84423.1; -  
DR InterPro: IPR004044; KH\_TYPE\_2;  
DR InterPro: IPR001351; Ribosomal\_S3.  
DR Pfam: PF00189; Ribosomal\_S3\_C; 1.  
DR Pfam: PF00417; Ribosomal\_S3\_N; 1.  
DR TIGRFAMs: TIGR01009; rpsC\_bact; 1.  
DR PROSITE: PS50823; KH\_TYPE\_2; 1.  
DR PROSITE: PS00548; RIBOSOMAL\_S3; 1.  
DR Ribosomal protein; Chloroplast.  
DOMAIN 47 118 KH TYPE-2.  
SEQUENCE 218 AA; 25188 MW; 763CB03FA6D74888 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 218;  
Best Local Similarity 83.3%; Pred. No. 60;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWI 6  
1:||||  
DB 209 IQIWI 214

RR3\_SPIOL STANDARD: PRT: 218 AA.

ID P09595;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Chloroplast 30S ribosomal protein S3.  
GN RPS3.  
OS Spinacia oleracea (Spinach).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
NCBI\_TaxID=3562;  
[1]

## SEQUENCE FROM N.A.

RC TISSUE=Leaf;  
RX MEDLINE=89313684; PubMed=2747623;  
RA Zhou D.X., Quigley F., Massenet O., Maché R.;  
RT "Cortranscription of the S10- and spc-like operons in spinach  
chloroplasts and identification of three of their gene products";  
RL Mol. Gen. Genet. 216:439-445(1989).  
[2]

## SEQUENCE FROM N.A.

RC STRAIN=cv. Geant d'hiver, and cv. Monatol;  
RX MEDLINE=21187424; PubMed=11292076;  
RA Schmitz-Lineweber C., Maier R.M., Alcatraz J.-P., Cottet A.,  
Herrmann R.G., Maché R.;  
RT "The plastid chromosome of spinach (Spinacia oleracea): complete  
nucleotide sequence and gene organization";  
RL Plant Mol. Biol. 45:307-315(2001).

## -1- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.

CC -1- SIMILARITY: CONTAINS 1 KH TYPE-2 DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL: X13336; CAA31715.1; -  
DR EMBL: AJ400848; CAB88766.1; -  
DR PIR: S01978; R3SP3.  
DR InterPro: IPR004044; KH\_TYPE\_2;  
DR InterPro: IPR001351; Ribosomal\_S3.  
DR Pfam: PF00189; Ribosomal\_S3\_C; 1.  
DR Pfam: PF00417; Ribosomal\_S3\_N; 1.  
DR TIGRFAMs: TIGR01009; rpsC\_bact; 1.  
DR PROSITE: PS50823; KH\_TYPE\_2; 1.  
DR PROSITE: PS00548; RIBOSOMAL\_S3; 1.  
DR Ribosomal protein; Chloroplast.  
DOMAIN 47 118 KH TYPE-2.  
SEQUENCE 218 AA; 24928 MW; 2D2FA40C67431000 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 218;  
Best Local Similarity 83.3%; Pred. No. 60;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWI 6  
1:||||  
DB 209 IQIWI 214

RR3\_TOBAC STANDARD: PRT: 218 AA.

ID P06357;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Chloroplast 30S ribosomal protein S3.  
GN RPS3.  
OS Nicotiana tabacum (Common tobacco).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
NCBI\_TaxID=4097;  
[1]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. Bright yellow 4;  
RA Shinozaki K., Ohme M., Tanaka M., Wakasugi T., Hayashida N.,

RA Matsubayashi T., Zaita N., Chunwongse J., Obokata J.,  
 RA Yamaguchi-Shinozaki K., Ohno C., Torazawa K., Meng B.-Y., Sugita M.,  
 RA Domo H., Kamogashira T., Yamada K., Kusuda J., Takaiwa F., Kato A.,  
 RA Tohdo N., Shimada H., Sugita M.,  
 RT "The complete nucleotide sequence of the tobacco chloroplast genome:  
 RT its gene organization and expression.",  
 RL EMBL J. 5:2043-2049(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86287388; PubMed=3016736;  
 RA Tanaka M., Wakasugi T., Sugita M., Shinozaki K., Sugita M.,  
 RT "Genes for the eight ribosomal proteins are clustered on the  
 RT chloroplast genome of tobacco (Nicotiana tabacum): similarity to the  
 RT S10 and spc operons of Escherichia coli.",  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:6030-6034(1986).  
 CC -1- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -1- SIMILARITY: CONTAINS 1 KH TYPE-2 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: 200044; CAA77381.1; -  
 DR PIR: A02702; R3NT3.  
 DR InterPro: IPR004044; KH\_TYPE-2.  
 DR InterPro: IPR001351; Ribosomal\_S3.  
 DR Pfam: PF00189; Ribosomal\_S3\_C; 1.  
 DR Pfam: PF00417; Ribosomal\_S3\_N; 1.  
 DR TIGRFAMS: TIGR01009; rpsc\_bact; 1.  
 DR PROSITE: PS50823; KH\_TYPE-2; 1.  
 DR PROSITE: PS00548; RIBOSOMAL\_S3; 1.  
 KW Ribosomal protein; Chloroplast.  
 FT DOMAIN 47 118 KH\_TYPE-2.  
 SQ SEQUENCE 218 AA; 25085 MW; 1750557D37B0D300 CRC64;  
 Query Match 88.2%; Score 30; DB 1; Length 218;  
 Best Local Similarity 83.3%; Pred. No. 60;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 IOIWF 6  
 Db 209 IKIWF 214  
 ID R3\_EPIVI STANDARD; PRT; 220 AA.  
 AC P30055;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Chloroplast 30S ribosomal protein S3.  
 GN RPS3.  
 OS Epifagus virginiana (Beechdrops).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; eunasterids I; Lamiales; Orchnanchaceae; Epifagus.  
 OX NCBI\_TaxID=41177;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93066301; PubMed=1332054;  
 RA Wolfe K.H., Morden C.W., Palmer J.D.;  
 RT "Function and evolution of a minimal plastid genome from a  
 RT nonphotosynthetic parasitic plant".  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:10648-10652(1992).  
 CC [2]  
 DR SEQUENCE FROM N.A.  
 RX MEDLINE=93021155; PubMed=1404416;

RA Wolfe K.H., Morden C.W., Ems S.C., Palmer J.D.;  
 RT "Rapid evolution of the plastid translational apparatus in a  
 RT nonphotosynthetic plant: loss or accelerated sequence evolution of  
 RT tRNA and ribosomal protein genes".  
 RL J. Mol. Evol. 35:304-317(1992).  
 CC -1- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -1- SIMILARITY: CONTAINS 1 KH TYPE-2 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M8184; AA65864.1; -  
 DR InterPro: IPR001351; Ribosomal\_S3.  
 DR Pfam: PF00189; Ribosomal\_S3\_C; 1.  
 DR Pfam: PF00417; Ribosomal\_S3\_N; 1.  
 DR TIGRFAMS: TIGR01009; rpsc\_bact; 1.  
 DR PROSITE: PS50823; KH\_TYPE-2; FALSE\_NEG.  
 DR PROSITE: PS00548; RIBOSOMAL\_S3; 1.  
 KW Ribosomal protein; Chloroplast.  
 FT DOMAIN 39 120 KH\_TYPE-2.  
 SQ SEQUENCE 220 AA; 25939 MW; 5BF70AC74AB7DF94 CRC64;  
 Query Match 88.2%; Score 30; DB 1; Length 220;  
 Best Local Similarity 83.3%; Pred. No. 61;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 IOIWF 6  
 Db 211 IKIWF 216  
 ID RT03\_CHOCR STANDARD; PRT; 240 AA.  
 AC P48938;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Mitochondrial ribosomal protein S3.  
 GN RPS3.  
 OS Chondrus crispus (Carrageen).  
 OG Mitochondrion.  
 OC Eukaryota; Rhodophyta; Florideophyceae; Gigartinales; Gigartinales;  
 OC Chondrus.  
 OX NCBI\_TaxID=2769;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE-APICES;  
 RX MEDLINE=95341681; PubMed=7616569;  
 RA Leblanc C., Boyen C., Richard O., Bonnard G., Grienenberger J.M.,  
 RA Kloareg B.;  
 RT "Complete sequence of the mitochondrial DNA of the rhodophyte  
 RT Chondrus crispus (Gigartinales). Gene content and genome  
 RT organization".  
 RL J. Mol. Biol. 250:484-495(1995).  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -1- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: 247547; CAA87601.1; -  
 DR InterPro: IPR001351; Ribosomal\_S3.

DR Pfam: PF00189; Ribosomal\_S3\_C; 1.  
 DR Pfam: PF00417; Ribosomal\_S3\_N; 1.  
 DR PROSITE: PS00548; RIBOSOMAL\_S3; FALSE\_NEG.  
 KW Ribosomal protein; Mitochondrion.  
 SO SEQUENCE 240 AA; 28142 MW; C910AA2E6D8C80DE CRC64;

Query Match 88.2%; Score 30; DB 1; Length 240;  
 Best Local Similarity 50.0%; Pred. No. 66;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 QIWIIF 6  
 DB 220 QIWIIF 225

RESULT 7  
 YL11\_CAEEL STANDARD; PRT; 277 AA.

AC Q11098;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 15-JUN-2002 (Rel. 41, Last annotation update)  
 15-JUN-2002 (Rel. 41, Last annotation update)  
 De Hypothetical protein C02P12.1 in chromosome X.  
 GN C02P12.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Palodoridae; Caenorhabditis.  
 NC NCBL\_TaxID=6239;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;  
 RL Miller N.;  
 RN Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

RE REVISIONS.

RA Waterston R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: SOME, TO THE TETRASPANIN (TM4SF) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: G41545; AAK39137.1; -  
 CC Wormpep: C02P12.1; CE25749.  
 CC InterPro: IPR000301; Transmem4.

DR Pfam: PF00335; transmembrane4; 1.  
 DR PROSITE: PS00421; TM4\_1; FALSE\_NEG.  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 29 49 POTENTIAL.  
 FT TRANSMEM 54 74 POTENTIAL.  
 FT TRANSMEM 239 259 POTENTIAL.  
 SO SEQUENCE 277 AA; 31753 MW; 8084221C8D4C541D CRC64;

Query Match 88.2%; Score 30; DB 1; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QIWIIF 6  
 DB 238 QIWIIF 242

RESULT 8  
 T2G1\_HAEGA STANDARD; PRT; 488 AA.  
 AC P43418;  
 DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Type II restriction enzyme HgaI (EC 3.1.21.4) (Endonuclease HgaI)  
 DE (R.HgaI).

GN Haemophilus gallinarum.  
 OS Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBL\_TaxID=728;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NCTC 3438;  
 RA Sugisaki H.;

RT "Nucleotide sequence of the gene of HgaI restriction endonuclease";  
 RL Bull. Inst. Chem. Res., Kyoto Univ. 71:338-342(1993).

CC -1- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCES GACGC AND GCGTC  
 CC AND CLEAVES RESPECTIVELY 10 BASES AFTER G-1 AND 10 BASES BEFORE  
 CC G'-1.  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give  
 CC specific double-stranded fragments with terminal 5'-phosphates.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: D17388; BAA04208.1; -  
 DR REBASE: 1096; HgaI.  
 KW Hydrolyase; Endonuclease; Nuclease; Restriction system.

SO SEQUENCE 488 AA; 56737 MW; B5E05259E3C710D2 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 488;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QIWIIF 6  
 DB 11 QIWIIF 15

RESULT 9  
 Y052\_BUCAI STANDARD; PRT; 144 AA.

AC P57160;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein BU052.  
 GN BU052.

OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum  
 OS symbiotic bacterium).  
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.

NC NCBL\_TaxID=118099;

RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-Tokyo 1998;  
 MEDLINE=20445173; PubMed=10993077;

RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;

RT "Genome sequence of the endocellular bacterial symbiont of aphids  
 RT Buchnera sp. APS";

RL Nature 407:81-86(2000).  
 CC -1- SIMILARITY: STRONG, TO E. COLI Y1B AND H. INFLUENZAE HI0744.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



CC -----  
 DR EMBL: AP001118; BAB12775.1; -  
 DR InterPro: IPR001763; Rhodanese-like.  
 DR Pfam: PF00581; Rhodanese; 1.  
 DR SMART: SM00450; RHOD; 1.  
 KM Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 144 AA; 16921 MW; 87D669CDA099ECB CRC64;

Query Match 85.3%; Score 29; DB 1; Length 144;  
 Best Local Similarity 83.3%; Pred. No. 61;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 IQIWIIF 6  
 I:::IIII  
 Db 15 ISIWIF 20

RESULT 10

RR3\_MESVI STANDARD; PRT; 213 AA.

09MU02;

16-OCT-2001 (Rel. 40, Created)

16-OCT-2001 (Rel. 40, Last sequence update)

15-JUN-2002 (Rel. 41, Last annotation update)

Chloroplast 30S ribosomal protein S3.

PS3.

Mesostigma viride.

OS Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;

OC Mesostigmatales; Mesostigmataceae; Mesostigma.

OX NCBI\_TaxID=41882;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN-NIES-296;

MEDLINE-20150907; PubMed-1068199;

Lemaux C., Otis C., Turmel M.;

"Ancstral chloroplast genome in Mesostigma viride reveals an early

branch of green plant evolution.";

Nature 403:649-652(2000).

-1- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.

-1- SIMILARITY: CONTAINS 1 KH TYPE-2 DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AF16114; AAF3809.1; -

DR InterPro: IPR004044; KH\_TYPE\_2.

DR InterPro: IPR004087; KH\_dom.

DR InterPro: IPR001351; Ribosomal\_S3.

DR Pfam: PF00013; KH-domain; 1.

DR Pfam: PF00189; Ribosomal\_S3\_C; 1.

DR Pfam: PF00417; Ribosomal\_S3\_N; 1.

DR SMART: SM00322; KH; 1.

DR TIGRfam: TIGR01009; rpsC\_bact; 1.

DR PROSITE: PS50823; KH\_TYPE\_2; 1.

DR PROSITE: PS00548; RIBOSOMAL\_S3; 1.

KW Ribosomal protein; Chloroplast.

FT DOMAIN 39 109 KH TYPE-2.

SQ SEQUENCE 213 AA; 24545 MW; 1C7ECC9A7B847CD CRC64;

Query Match 85.3%; Score 29; DB 1; Length 213;

Best Local Similarity 66.7%; Pred. No. 90;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IQIWIIF 6  
 I:::IIII  
 Db 200 IKWIF 205

RESULT 11

RR3\_MARPO

ID RR3\_MARPO STANDARD; PRT; 217 AA.

AC P06356;

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Chloroplast 30S ribosomal protein S3.

GN RPS3.

OS Marchantia polymorpha (liverwort).

OG Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;

OC Marchantiopsida; Marchantiidae; Marchantiales; Marchantiaceae;

OC Marchantiaceae; Marchantia.

OX NCBI\_TaxID=3197;

RN [1]

SEQUENCE FROM N.A.

RX MEDLINE-89068687; PubMed-3199436;

RA Fukuzawa H., Kohchi T., Sano T., Shirai H., Umesono K., Inokuchi H.,

Ozeki H., Ohyama K.,

"Structure and organization of Marchantia polymorpha chloroplast

genome. III. Gene organization of the large single copy region from

rbcl to trnI(CAU).";

J. Mol. Biol. 203:333-351(1988).

[2]

SEQUENCE FROM N.A.

RA Ohyama K., Fukuzawa H., Kohchi T., Shirai H., Sano T., Sano S.,

Umesono K., Shiki Y., Takeuchi M., Chang Z., Aota S., Inokuchi H.,

Ozeki H.;

"Chloroplast gene organization deduced from complete sequence of

Nature 322:572-574(1986).

-1- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.

-1- SIMILARITY: CONTAINS 1 KH TYPE-2 DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: X04465; CAA28124.1; -

DR PIR: A02701; R3LV3.

DR InterPro: IPR004044; KH\_TYPE\_2.

DR InterPro: IPR001351; Ribosomal\_S3.

DR Pfam: PF00189; Ribosomal\_S3\_C; 1.

DR Pfam: PF00417; Ribosomal\_S3\_N; 1.

DR TIGRfam: TIGR01009; rpsC\_bact; 1.

DR PROSITE: PS50823; KH\_TYPE\_2; 1.

DR PROSITE: PS00548; RIBOSOMAL\_S3; 1.

KW Ribosomal protein; Chloroplast.

FT DOMAIN 46 117 KH TYPE-2.

SQ SEQUENCE 217 AA; 25055 MW; 7AF3CA7834B8A01 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 217;

Best Local Similarity 66.7%; Pred. No. 92;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IQIWIIF 6  
 I:::IIII  
 Db 208 IKWIF 213

RESULT 12

RR3\_MAIZE

ID RR3\_MAIZE STANDARD; PRT; 224 AA.

AC P06586;

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

DE Chloroplast 30S ribosomal protein S3.
GN RPS3.
OS Zea mays (Maize).
OC Chloroplast.
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
CC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87231045; PubMed=3588305;
RA McLaughlin W.E., Larrina I.M.;
RT "The sequence of the maize plastid encoded rps3 locus.";
RL Nucleic Acids Res. 15:4689-4689(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95395841; PubMed=7666415;
RA Mäler R.M., Neckermann K., Igloi G.L., Koessel H.;
RT "Complete sequence of the maize chloroplast genome: gene content,
RT hotspots of divergence and fine tuning of genetic information by
RT transcript editing.";
CC J. Mol. Biol. 251:614-628(1995).
CC -1- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.
CC -1- SIMILARITY: CONTAINS 1 KH TYPE-2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00340; CAA68427.1; -
DR EMBL: X86563; CAA60324.1; -
DR MaizeDB: 66303; -
DR InterPro: IPR001351; Ribosomal_S3.
DR Pfam: PF00189; Ribosomal_S3_C; 1.
DR Pfam: PF00417; Ribosomal_S3_N; 1.
DR TIGRFAMs: TIGR01009; rpsC_bact; 1.
DR PROSITE: PS50823; KH_TYPE_2; FALSE_NEG.
DR PROSITE: PS00548; RIBOSOMAL_S3; 1.
KW Ribosomal protein; Chloroplast.
FT DOMAIN 43 124 KH TYPE-2.
SQ SEQUENCE 224 AA; 25916 MW; B6B8B6BA8AF70DA CRC64;

Query Match
Best Local Similarity 85.3%; Score 29; DB 1; Length 224;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 IQIWF 6
:::||||
DB 215 VKIWF 220

RESULT 13
RR3_CHLVU STANDARD: PRT; 231 AA.
AC P56365;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S3.
GN RPS3.
OS Chlorella vulgaris.
OC Chloroplast.
CC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
CC Chlorellaceae; Chlorella.
OX NCBI_TaxID=3077;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=IAM C-27 / Tam1ya;
RX MEDLINE=97303241; PubMed=9159184;

```

```

RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,
RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,
RA Inamura A., Yoshinaga K., Sugita M.;
RT "Complete nucleotide sequence of the chloroplast genome from the
RT green alga Chlorella vulgaris: the existence of genes possibly
RT involved in chloroplast division.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).
CC -1- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.
CC -1- SIMILARITY: CONTAINS 1 KH TYPE-2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB001684; BAA58006.1; -
DR InterPro: IPR004044; KH_TYPE_2.
DR InterPro: IPR001351; Ribosomal_S3.
DR Pfam: PF00189; Ribosomal_S3_C; 1.
DR Pfam: PF00417; Ribosomal_S3_N; 1.
DR TIGRFAMs: TIGR01009; rpsC_bact; 1.
DR PROSITE: PS50823; KH_TYPE_2; 1.
DR PROSITE: PS00548; RIBOSOMAL_S3; 1.
KW Ribosomal protein; Chloroplast.
FT DOMAIN 39 123 KH TYPE-2.
SQ SEQUENCE 231 AA; 26404 MW; D3EDE90A93551097 CRC64;

Query Match
Best Local Similarity 85.3%; Score 29; DB 1; Length 231;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 IQIWF 6
1:11:1
DB 214 IKIWF 219

RESULT 14
RR3_ORYSA STANDARD: PRT; 239 AA.
AC P12146;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S3.
GN RPS3.
OS Oryza sativa (Rice).
OC Chloroplast.
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
CC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Nipponbare;
RX MEDLINE=89364698; PubMed=2770692;
RA Hiratsuka J., Shimada H., Whittier R., Ishibashi T., Sakamoto M.,
RA Mori M., Kondo C., Honji Y., Sun C.-R., Meng B.-Y., Li Y.-Q.,
RA Kanno A., Nishizawa Y., Hirai A., Shinozaki K., Sugita M.;
RT "The complete sequence of the rice (Oryza sativa) chloroplast genome:
RT intermolecular recombination between distinct tRNA genes accounts for
RT a major plastid DNA inversion during the evolution of the cereals.";
RL Mol. Gen. Genet. 217:185-194(1989).
CC -1- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.
CC -1- SIMILARITY: CONTAINS 1 KH TYPE-2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X15901; CAA33934.1; 1.
DR  PIR; J00265; R3R23.
DR  InterPro; IPR001351; Ribosomal_S3.
DR  Pfam; PF00189; Ribosomal_S3_C; 1.
DR  Pfam; PF00417; Ribosomal_S3_N; 1.
DR  TIGRFAMs; TIGR01009; rpsc_bact; 1.
DR  PROSITE; PS50823; KH_type_2; FALSE_NEG.
DR  Ribosomal protein; CHLORoplast.
KW  DOMAIN
SQ  SEQUENCE 239 AA; 27518 MW; 3152AC028BFCC139 CRC64;

Query Match
Best Local Similarity 85.3%; Score 29; DB 1; Length 239;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 IQIWIIF 6
Db 230 VKIWIIF 235
```

```
RESULT 15
RS3_SYNY3
ID RS3_SYNY3 STANDARD; PRT; 239 AA.
AC P73314;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S3.
GN RPS3 OR RPS3 OR SLI804.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxId=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naito K.,
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RT DNA Res. 3:109-136(1996).
CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE BINDING OF INITIATOR
CC MET-TRNA (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE S3P FAMILY OF RIBOSOMAL PROTEINS.
CC -1- SIMILARITY: CONTAINS 1 KH TYPE-2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90905; BAA17343.1; 1.
CC InterPro; IPR004044; KH_type_2.
CC InterPro; IPR004087; KH_dom.
CC InterPro; IPR004086; KH_type_1.
CC InterPro; IPR001351; Ribosomal_S3.
CC Pfam; PF00189; Ribosomal_S3_C; 1.
CC Pfam; PF00417; Ribosomal_S3_N; 1.
CC SMART; SM00322; KH; 1.
CC TIGRFAMs; TIGR01009; rpsc_bact; 1.
CC PROSITE; PS50823; KH_type_2; 1.
CC PROSITE; PS00548; RIBOSOMAL_S3; 1.
```

```
KW  Ribosomal protein; tRNA-binding; Complete proteome.
FT  INIT_MET 0
FT  DOMAIN 64 100 KH TYPE-2.
SQ  SEQUENCE 239 AA; 27016 MW; 066ACE072B38C6B2 CRC64;

Query Match
Best Local Similarity 85.3%; Score 29; DB 1; Length 239;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

1 IQIWIIF 6
Db 199 IKWIIF 204

Search completed: January 3, 2003, 15:29:22
Job time : 6.18182 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:24:35 ; Search time 20.3182 Seconds  
(without alignments)  
60.846 Million cell updates/sec

Title: US-09-543-188a-21

Perfect score: 34  
Sequence: 1 IQIWI 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

al number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp-archaea:\*
- 2: sp-bacteria:\*
- 3: sp-fungi:\*
- 4: sp-human:\*
- 5: sp-invertebrate:\*
- 6: sp-mammal:\*
- 7: sp-mhc:\*
- 8: sp-organellar:\*
- 9: sp-phage:\*
- 10: sp-plant:\*
- 11: sp-rodent:\*
- 12: sp-virus:\*
- 13: sp-vertebrate:\*
- 14: sp-unclassified:\*
- 15: sp-virus:\*
- 16: sp-bacteriap:\*
- 17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	97.1	282	8 Q9Z2N7	Q9Z2N7 cyanidiosch
2	33	97.1	285	5 O17997	O17997 caenorhabd
3	33	97.1	360	16 Q9KCM4	Q9KCM4 bacillus ha
4	32	94.1	86	10 Q9MOC7	Q9MOC7 arabidopsis
5	32	94.1	88	10 Q9ZV45	Q9ZV45 arabidopsis
6	32	94.1	291	16 Q9CLF4	Q9CLF4 pasteurella
7	32	94.1	495	8 Q63853	Q63853 sarcophyton
8	31	91.2	320	2 Q93HD9	Q93HD9 streptomyce
9	31	91.2	394	16 Q9PMP4	Q9PMP4 campylobact
10	31	91.2	719	2 Q8RPJ4	Q8RPJ4 desulfatoba
11	30	88.2	115	11 Q9CU03	Q9CU03 mus musculu
12	30	88.2	218	8 Q9MT17	Q9MT17 oenothera h
13	30	88.2	218	10 Q8S8V5	Q8S8V5 atropa bell
14	30	88.2	229	16 Q66753	Q66753 aquilex aeo
15	30	88.2	246	3 Q9P3R7	Q9P3R7 neurospora
16	30	88.2	320	2 Q9RNE1	Q9RNE1 bacillus an

17	30	88.2	344	17 Q973N4	Q973N4 sulfolobus
18	30	88.2	364	2 Q85468	Q85468 bacillus ce
19	30	88.2	403	17 Q9HLS6	Q9HLS6 thermoplas
20	30	88.2	427	17 Q97C83	Q97C83 thermoplas
21	30	88.2	563	10 Q8W0K1	Q8W0K1 oryza sativ
22	29	85.3	107	5 Q61555	Q61555 drosophila
23	29	85.3	164	11 Q8R460	Q8R460 mus musculu
24	29	85.3	218	8 Q98455	Q98455 spirogyra m
25	29	85.3	218	8 Q9BBP8	Q9BBP8 lotus japon
26	29	85.3	219	16 Q8RIG1	Q8RIG1 fusobacteri
27	29	85.3	221	8 Q9T120	Q9T120 nephrusem1
28	29	85.3	222	16 Q97P66	Q97P66 streptococc
29	29	85.3	235	16 Q9CL37	Q9CL37 pasteurella
30	29	85.3	239	8 Q95H49	Q95H49 triticum ae
31	29	85.3	243	16 Q8UE24	Q8UE24 agrobacteri
32	29	85.3	244	8 Q99992	Q99992 porphyra pu
33	29	85.3	252	2 Q8V0R8	Q8V0R8 phytoplasm
34	29	85.3	255	11 Q9C2V2	Q9C2V2 mus musculu
35	29	85.3	327	16 P74743	P74743 synchocyst
36	29	85.3	348	16 Q51706	Q51706 borrelia bu
37	29	85.3	373	5 Q61643	Q61643 drosophila
38	29	85.3	417	3 Q9C473	Q9C473 emericella
39	29	85.3	523	8 Q9G8Q4	Q9G8Q4 naegleria g
40	29	85.3	553	10 Q64861	Q64861 arabidopsis
41	29	85.3	616	17 Q9HIF5	Q9HIF5 thermoplas
42	29	85.3	678	5 Q8WR60	Q8WR60 drosophila
43	29	85.3	720	10 Q9SPF9	Q9SPF9 arabidopsis
44	29	85.3	720	10 Q9SPF8	Q9SPF8 arabidopsis
45	29	85.3	720	10 Q80809	Q80809 arabidopsis

#### ALIGNMENTS

RESULT 1

ID	Q9Z2N7	PRELIMINARY;	PRT;	282 AA.
AC	Q9Z2N7			
DT	01-MAY-1999 (T-EMBLrel. 10, Created)			
DT	01-MAY-1999 (T-EMBLrel. 10, Last sequence update)			
DE	01-MAR-2002 (T-EMBLrel. 20, Last annotation update)			
DE	30S ribosomal protein S3.			
GN	RPS3.			
OS	Cyanidioschyzon merolae.			
OG	Mitochondrion.			
OC	Eukaryota; Rhodophyta; Bangiophyceae; Cyanidioschyzon.			
OX	NCBI_TaxID=45157;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=10D;			
RX	MEDLINE=99030526; PubMed=9801318;			
RA	Ohta N., Sato N., Kuroiwa T.;			
RT	"Structure and organization of the mitochondrial genome of the			
RT	unicellular red alga Cyanidioschyzon merolae deduced from the complete			
RT	nucleotide sequence."			
RL	Nucleic Acids Res. 26:5190-5198(1998).			
DR	EMBL: D89861; BAA36537.1.-.			
DR	InterPro: IPR001351; Ribosomal_S3.			
DR	Pfam: PF00189; Ribosomal_S3_C; 1.			
KW	Mitochondrion; Ribosomal protein.			
SQ	SEQUENCE 282 AA; 33871 MW; 89CC910DAA22DFB2 CRC64;			

Query Match 97.1%; Score 33; DB 8; Length 282;  
Best local Similarity 83.3%; Pred. No. 1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWI 6  
Db 275 IQYWI 280  
RESULT 2  
O17997

ID 017997 PRELIMINARY: PRT: 285 AA.  
AC 017997;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE R10012.11 protein.  
GN R10012.11.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RL Percy C.M.;  
RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
investigating biology."  
Science 282:2012-2018(1998).  
DR EMBL; Z61109; CAB03254.1; -  
DR InterPro: IPR003003; 7TM\_chemo2.  
DR InterPro: IPR00168; 7TM\_nematode.  
DR Pfam: PF01604; 7tm\_5; 1.  
SQ SEQUENCE 285 AA; 32927 MW; C3078F66F3561CBD CRC64;

Query Match 97.1%; Score 33; DB 5; Length 285;  
Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IQIWF 6  
:||||  
DB 48 VOIWF 53

RESULT 3  
O9KCM4 PRELIMINARY: PRT: 360 AA.  
ID 09KCM4;  
AC 09KCM4;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE Spore germination protein.  
GN BH1455  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Bacillaceae; Bacillus.  
NX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirana C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
halodurans and genomic sequence comparison with Bacillus subtilis";  
Nucleic Acids Res. 28:4317-4331(2000).  
DR EMBL; AP001512; BAB05174.1; -  
KN Complete proteome.  
SQ SEQUENCE 360 AA; 41587 MW; 1B25DC9DAF6E328 CRC64;

Query Match 97.1%; Score 33; DB 16; Length 360;  
Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IQIWF 6  
:||||  
DB 109 IOIWF 114

RESULT 4  
O9MOC7 PRELIMINARY: PRT: 86 AA.  
ID 09MOC7;  
AC 09MOC7;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE Small nuclear ribonucleoprotein homolog.  
GN AT4G30330.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucots II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AL161576; CAB81026.1; -  
DR InterPro: IPR001163; snRNP\_Sm.  
DR Pfam: PF01423; Sm; 1.  
KW Nucleocapsid; Ribonucleoprotein.  
SQ SEQUENCE 86 AA; 10097 MW; E5AD9FDF35C3AC0D CRC64;

Query Match 94.1%; Score 32; DB 10; Length 86;  
Best Local Similarity 83.3%; Pred. No. 50;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IQIWF 6  
:||||  
DB 27 IOIWF 32

RESULT 5  
O9ZV45 PRELIMINARY: PRT: 88 AA.  
ID 09ZV45;  
AC 09ZV45;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Putative small nuclear ribonucleoprotein E.  
GN AT2G18740.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucots II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RX MEDLINE=20083487; PubMed=10617197;  
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
RA Buehl C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Motilal K.S.,  
RA Adams L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,  
RA Cronin M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
RA Copenhaver G.P., Preuss D., Nielsen W.C., White O., Eisen J.A.,  
RA Salzberg S.L., Fraser C.M., Venter J.C.;  
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
thaliana";  
RL Nature 402:761-768(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Lin X.;  
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AC005724; AAD08943.1; -  
DR InterPro: IPR001163; snRNP\_Sm.

DR Pfam: PF01423; Sm; 1.  
 KW Nucleocapsid; Ribonucleoprotein.  
 SQ SEQUENCE 88 AA; 10298 MW; 762719DD9FDF35C5 CRC64;

Query Match 94.1%; Score 32; DB 10; Length 88;  
 Best Local Similarity 83.3%; Pred. No. 52;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IQIWF 6  
 |||||:  
 Db 27 IQIWLIF 32

RESULT 6  
 O9CLF4

ID O9CLF4 PRELIMINARY; PRT; 291 AA.  
 AC O9CLF4;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 RT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 GN Hypothetical protein PM1281.  
 OS PM1281.  
 OC Pasteurella multocida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Pasteurella.  
 OX NCBI\_TaxID=747;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=PM70;  
 RX MEDLINE=21145866; PubMed=11248100;  
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
 RT "Complete genomic sequence of Pasteurella multocida PM70.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
 DR EMBL: AF006167; AAK03365.1; -;  
 DR InterPro: IPR004626; RARD; 1.  
 DR TrEMBL: TIGR00688; rard; 1.  
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 291 AA; 32530 MW; 43F47FDE6986512B CRC64;

Query Match 94.1%; Score 32; DB 16; Length 291;  
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IQIWF 6  
 |||||:  
 Db 84 IQIWLIF 89

RESULT 7  
 O63853

ID O63853 PRELIMINARY; PRT; 495 AA.  
 AC O63853;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 RT 01-AUG-1999 (TREMBLrel. 11, Last sequence update)  
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 GN NADH dehydrogenase subunit 4.  
 OS NDA.  
 OC Sarcophyton glaucum.  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Alcyonaria; Alcyonacea;  
 OC Alcyonidae; Sarcophyton.  
 OX NCBI\_TaxID=70919;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA MEDLINE=99065763; PubMed=9847412;  
 RA Beaton M.J., Roger A.J., Cavalier-Smith T.;  
 RT "Sequence analysis of the mitochondrial genome of Sarcophyton glaucum:  
 RT conserved gene order among octocorals.";  
 RL J. Mol. Evol. 47:697-708(1998).  
 DE [2]  
 GN SEQUENCE OF 450-495 FROM N.A.

RP MEDLINE=98210232; PubMed=9541536;  
 RX Pont-Kington G., Okada N.A., Macfarlane J.L., Beagley C.T.,

RA Watkins-Sims C.D., Cavalier-Smith T., Clark-Walker G.D.,  
 RA Moltenhime D.R.;  
 RT "Mitochondrial DNA of the coral Sarcophyton glaucum contains a gene  
 RT for a homologue of bacterial Puts: a possible case of gene transfer  
 RT from the nucleus to the mitochondrion.";  
 RL J. Mol. Evol. 46:419-431(1998).  
 CC -1-CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.

DR EMBL: AF064823; AAC99651.1; -;  
 DR EMBL: AF063192; AAC16387.1; -;  
 DR InterPro: IPR001750; Oxidored\_q1.  
 DR Pfam: PF00361; oxidored\_q1; 1.  
 DR PRINTS: PR01559; DUFFYANTIGEN.  
 KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.  
 SQ SEQUENCE 495 AA; 55419 MW; 6E0C04C478A5CCB4 CRC64;

Query Match 94.1%; Score 32; DB 8; Length 495;  
 Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IQIWF 6  
 |||||:  
 Db 219 IQIWLIF 224

RESULT 8  
 O93HD9

ID O93HD9 PRELIMINARY; PRT; 320 AA.  
 AC O93HD9;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 RT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 GN ABC transporter.  
 OS Streptomyces avermitilis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=33903;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=21477403; PubMed=11572948;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinohe M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: Deducing the ability of producing secondary  
 RT metabolites.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 DR EMBL: AB070945; BAB69249.1; -;  
 DR InterPro: IPR000515; BPD\_transp.  
 DR Pfam: PF00528; BPD\_transp; 1.  
 DR PROSITE: PS00402; BPD\_TRANSF\_INN\_MEMBER; UNKNOWN.1.

SQ SEQUENCE 320 AA; 34176 MW; DD3BC5AB26756426 CRC64;

Query Match 91.2%; Score 31; DB 2; Length 320;  
 Best Local Similarity 66.7%; Pred. No. 2.8e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IQIWF 6  
 :|||:  
 Db 308 VOLWIF 313

RESULT 9  
 O9PMP4

ID O9PMP4 PRELIMINARY; PRT; 394 AA.  
 AC O9PMP4;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 RT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 GN Possible polysaccharide modification protein.  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
 OC Campylobacter.

OX NCBI\_TaxID=197;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCCTC 11168;  
RX MEDLINE=20150912; PubMed=10688204;  
RA Parthill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,  
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,  
RA Jagsals K., Kariyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
RA Whitehead S., Barrett B.G.;  
RT "The genome sequence of the food-borne pathogen *Campylobacter jejuni*  
RT reveals hypervariable sequences.";  
RL Nature 403:665-668(2000).  
DR EMBL: AL139078; CAB73837.1; -  
KM Complete proteome.  
SQ SEQUENCE 394 AA; 46938 MW; E4293DB5879FD64E CRC64;

Query Match 91.2%; Score 31; DB 16; Length 394;  
Best Local Similarity 83.3%; Pred. No. 3.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 105 IOWIF 110  
1 IOWIF 6  
1:11111

RESULT 10  
OBRP4 PRELIMINARY; PRT; 719 AA.  
AC OBRP4;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)  
DE 01-JUN-2002 (TREMBlrel. 21, last annotation update)  
OS Desulfitobacterium hafnense.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
OC Clostridiales; Peptococcaceae; Desulfitobacterium.  
OX NCBI\_TaxID=49338;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DCB-2;  
RA Davis J.K., Tiedje J.M.;  
RT "Sequence and transcriptional analysis of reductive dehalogenase genes  
RT of *Desulfitobacterium*."  
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF403182; AAL87758.1; -  
KM Hypothetical protein  
SQ SEQUENCE 719 AA; 80119 MW; B806F0F0FID79B CRC64;

Query Match 91.2%; Score 31; DB 2; Length 719;  
Best Local Similarity 66.7%; Pred. No. 6.2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IOWIF 6  
1:11111

Db 108 IOWIF 113

RESULT 11  
OQCU03 PRELIMINARY; PRT; 115 AA.  
AC OQCU03;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)  
DE 01-JUN-2001 (TREMBlrel. 17, last annotation update)  
DE 1700116B05R1k protein (fragment).  
GN 1700116B05R1k.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=TESTIS;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Aachii J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Knehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schmiel L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baladrelli R., Barsh G.,  
RA Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohatsu S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL: AK018963; BAB31497.1; -  
DR MGD: MGI:1925889; 1700116B05R1k.  
FT NON\_TER 1  
SQ SEQUENCE 115 AA; 12580 MW; F95076419A6A69C CRC64;

Query Match 88.2%; Score 30; DB 11; Length 115;  
Best Local Similarity 83.3%; Pred. No. 1.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IOWIF 6  
1:11111

Db 40 IOWIF 45

RESULT 12  
OQWT17 PRELIMINARY; PRT; 218 AA.  
AC OQWT17;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)  
DE 01-JUN-2002 (TREMBlrel. 21, last annotation update)  
DE Ribosomal protein S3.  
GN RPS3.  
OS Oenothera hookeri (Hooker's evening primrose).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Myrtales; Onagraceae; Oenothera.  
OX NCBI\_TaxID=85636;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20309318; PubMed=10852478;  
RA Hupfer H., Swiatek M., Hornung S., Herrmann R.G., Mäler R.M.,  
RA Chiu W.L., Sears B.;  
RT "Complete nucleotide sequence of the *Oenothera elata* plastid  
RT chromosome, representing plastome I of the five distinguishable  
RT *Euenothera* plastomes.";  
RL Mol. Gen. Genet. 263:581-585(2000).  
DR EMBL: AJ271079; CAB67198.1; -  
DR InterPro: IPR004044; KH\_TYPE\_2.  
DR Pfam: PF00189; Ribosomal\_S3\_C; 1.  
DR DR Pfam: PF00417; Ribosomal\_S3\_N; 1.  
DR TIGRFAMs: TIGR01009; tpsC\_bact; 1.  
DR PROSITE: PS50823; KH\_TYPE\_2; 1.  
DR PROSITE: PS00548; RIBOSOMAL\_S3; UNKNOWN\_1.  
KM Chloroplast.  
SQ SEQUENCE 218 AA; 24971 MW; B0B741F658B010BA CRC64;

Query Match 88.2%; Score 30; DB 8; Length 218;  
Best Local Similarity 83.3%; Pred. No. 3e+02;



Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IQIWF 6  
 1:||||  
 Db 210 IKIWF 215

## RESULT 13

OS8SV5 PRELIMINARY; PRT; 218 AA.  
 AC OS8SV5;  
 DT 01-JUN-2002 (TRENBLREL. 21, Created)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE Ribosomal protein S3.  
 GN RPS3.  
 OS Atropa belladonna (belladonna) (deadly nightshade).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Atropa.  
 NCBI\_TaxID=33113;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AB5P(KAN);  
 RA Schmitz-Linneweber C., Regel R., Gia Du T., Hupfer H., Herrmann R.G.,  
 RA Mater R.M.;  
 RT "The nucleotide sequence of the plastid chromosome of *Atropa*  
 RT *belladonna* (deadly nightshade) and its comparison with that of  
 RT *Nicotiana tabacum* with emphasis on sequence elements relevant for  
 RT microevolution."  
 RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AJ316582; CAC88082.1; -  
 SQ SEQUENCE 218 AA; 25042 MW; 115967F1CCED2F9 CRC64;

Query Match 88.2%; Score 30; DB 10; Length 218;  
 Best Local Similarity 83.3%; Pred. No. 3e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IQIWF 6  
 1:||||  
 Db 209 IKIWF 214

## RESULT 14

OS6753 PRELIMINARY; PRT; 229 AA.  
 AC OS6753;  
 DT 01-AUG-1998 (TRENBLREL. 07, Created)  
 DT 01-AUG-1998 (TRENBLREL. 07, Last sequence update)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
 DE Glutaredoxin-like protein.  
 GN GUA OR AQ.443.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.  
 NCBI\_TaxID=63363;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF5;  
 RX MEDLINE=98196666; PubMed=9537320;  
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay N., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium *Aquifex*  
 RT *aeolicus*."  
 RL Nature 392:353-358(1998).  
 DR EMBL: AE000690; AAC06709.1; -  
 DR InterPro: IPR000631; ThioRed.  
 DR InterPro: IPR000531; TonB\_boxC.  
 DR PROSITE: PS00430; TONB\_DEPENDENT\_REC\_1; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 229 AA; 25640 MW; C3A3C141FE05B25E CRC64;

Query Match 88.2%; Score 30; DB 16; Length 229;

Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IQIWF 6  
 1:||||  
 Db 138 IETWF 143

## RESULT 15

OS9P37 PRELIMINARY; PRT; 246 AA.  
 AC OS9P37;  
 DT 01-OCT-2000 (TRENBLREL. 15, Created)  
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE Related to cell protein.  
 GN B24P7.180.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Sordariaceae; Neurospora.  
 NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Schulte U., Aign V., Hohelsel J., Brandt P., Fartmann B., Holland R.,  
 RA Mykatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AL389890; CAB97283.2; -  
 DR InterPro: IPR005103; Glyco\_hydro\_61.  
 DR Pfam: PF03443; Glyco\_hydro\_61; 2.  
 SQ SEQUENCE 246 AA; 26817 MW; 30E4F3436791A92A CRC64;

Query Match 88.2%; Score 30; DB 3; Length 246;  
 Best Local Similarity 50.0%; Pred. No. 3.4e+02;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IQIWF 6  
 1:||||  
 Db 98 VOVWLF 103

Search completed: January 3, 2003, 15:32:02  
 Job time : 23.3182 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:16 : Search time 9 Seconds  
(without alignments)  
19.615 Million cell updates/sec

Title: US-09-543-188A-21

Perfect score: 34

Sequence: 1 IQIWF 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Indexed: 262574 seqs, 29422922 residues

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents-AA:  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	88.2	70	4	US-09-288-143-114
2	29	85.3	391	3	US-08-968-563-11
3	29	85.3	391	4	US-08-969-683A-11
4	29	85.3	391	4	US-09-297-928-7
5	29	85.3	433	2	US-08-919-624-1
6	28	82.4	68	4	US-09-025-151-18
7	28	82.4	287	3	US-09-065-474-145
8	28	82.4	287	4	US-09-557-034-145
9	28	82.4	312	2	US-09-031-485-2
10	28	82.4	312	2	US-08-847-429A-2
11	28	82.4	312	3	US-09-065-474-2
12	28	82.4	312	4	US-09-557-034-2
13	28	82.4	413	4	US-09-134-001C-5503
14	28	82.4	496	4	US-09-134-001C-3703
15	28	82.4	591	4	US-09-134-001C-3275
16	28	82.4	1745	2	US-09-031-485-33
17	28	82.4	1745	2	US-08-847-429A-33
18	28	82.4	1745	3	US-09-065-474-33
19	28	82.4	1745	4	US-09-557-034-33
20	27	79.4	15	1	US-08-148-160-2
21	27	79.4	51	1	US-07-791-213D-1
22	27	79.4	51	1	US-07-791-213D-2
23	27	79.4	51	1	US-07-791-213D-7
24	27	79.4	51	1	US-07-791-213D-17
25	27	79.4	51	1	US-07-972-387-67
26	27	79.4	51	1	US-08-431-412-67
27	27	79.4	51	1	US-08-057-971-67

28	27	79.4	51	1	US-08-293-150A-1	Sequence 1, Appl
29	27	79.4	51	1	US-08-293-150A-2	Sequence 2, Appl
30	27	79.4	51	1	US-08-293-150A-7	Sequence 7, Appl
31	27	79.4	51	1	US-08-293-150A-17	Sequence 17, Appl
32	27	79.4	54	1	US-07-791-213D-16	Sequence 16, Appl
33	27	79.4	54	1	US-08-293-150A-16	Sequence 16, Appl
34	27	79.4	58	1	US-07-700-526-11	Sequence 11, Appl
35	27	79.4	58	1	US-08-384-489-11	Sequence 11, Appl
36	27	79.4	58	1	US-08-358-160-3	Sequence 3, Appl
37	27	79.4	58	1	US-08-358-160-75	Sequence 75, Appl
38	27	79.4	58	1	US-08-463-155A-39	Sequence 39, Appl
39	27	79.4	58	1	US-08-463-432B-47	Sequence 39, Appl
40	27	79.4	58	1	US-08-676-125A-47	Sequence 47, Appl
41	27	79.4	58	1	US-08-206-310A-39	Sequence 39, Appl
42	27	79.4	58	1	US-08-398-010A-39	Sequence 39, Appl
43	27	79.4	58	2	US-08-398-628A-39	Sequence 39, Appl
44	27	79.4	58	2	US-08-399-115A-39	Sequence 39, Appl
45	27	79.4	58	2	US-09-136-012A-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1  
US-09-288-143-114  
; Sequence 114, Application US/09288143  
; Patent No. 6433139  
; GENERAL INFORMATION:  
; APPLICANT: Brewer et al.  
; TITLE OF INVENTION: 53 Human Secreted Proteins  
; FILE REFERENCE: P2018P1  
; CURRENT APPLICATION NUMBER: US/09/288,143  
; EARLIER APPLICATION NUMBER: PCT/US98/21142  
; EARLIER FILING DATE: 1998-10-08  
; EARLIER APPLICATION NUMBER: 60/061,463  
; EARLIER FILING DATE: 1997-10-09  
; EARLIER APPLICATION NUMBER: 60/061,529  
; EARLIER FILING DATE: 1997-10-09  
; EARLIER APPLICATION NUMBER: 60/071,498  
; EARLIER FILING DATE: 1997-10-09  
; EARLIER APPLICATION NUMBER: 60/061,527  
; EARLIER FILING DATE: 1997-10-09  
; EARLIER APPLICATION NUMBER: 60/061,536  
; EARLIER FILING DATE: 1997-10-09  
; EARLIER APPLICATION NUMBER: 60/061,532  
; EARLIER FILING DATE: 1997-10-09  
; NUMBER OF SEQ ID NOS: 219  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 114  
; LENGTH: 70  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (70)  
; OTHER INFORMATION: xaa equals stop translation  
US-09-288-143-114  
Query Match 88.2% Score 30; DB 4; Length 70;  
Best Local Similarity 50.0% Pred No. 1.2e+02;  
Matches 3; Conservative 3; Mismatches 0; Indels 0;  
Gaps 0;  
OY 1 IQIWF 6  
; : : : : :  
Db 15 LQWVF 20  
RESULT 2  
US-08-968-563-11  
; Sequence 11, Application US/08968563  
; Patent No. 6013494  
; GENERAL INFORMATION:

APPLICANT: CHARLES E. NAKAMURA  
APPLICANT: ANTHONY A. GATENBY  
APPLICANT: AMY (KUANG-HUA) HSU  
APPLICANT: RICHARD D. LA REAU  
APPLICANT: SHARON L. HAYNIE  
APPLICANT: MARIA DIAZ-FORRES  
APPLICANT: DONALD E. TRIMBUR  
APPLICANT: GREGORY M. WHITED  
APPLICANT: VASANTHA NAGARAJAN  
APPLICANT: MARK S. PAYNE  
APPLICANT: STEPHEN K. PICATAGGIO  
APPLICANT: RAMESH V. NAIR  
TITLE OF INVENTION: METHOD FOR THE RECOMBINANT  
TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANEDIOL  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY  
STREET: 1007 MARKET STREET  
CITY: WILMINGTON  
STATE: DELAWARE  
COUNTRY: U.S.A.  
ZIP: 19898  
ADDRESSEE: GENENCOR INTERNATIONAL, INC.  
STREET: 4 CAMBRIDGE PLACE  
STREET: 1870 SOUTH WINTON ROAD  
CITY: ROCHESTER  
STATE: NEW YORK  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.50 INCH DISKETTE  
OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95  
SOFTWARE: MICROSOFT WORD VERSION 7.0A  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/968,563  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/030,601  
FILING DATE: NOVEMBER 13, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: FLOYD, LINDA AXAMENHY  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: CR-9982  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-892-8112  
TELEFAX: 302-773-0164  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: GPD1  
US-08-968-563-11

Query Match 85.3%; Score 29; DB 3; Length 391;  
Best Local Similarity 50.0%; Pred. No. 8.2e+02;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 IQIWF 6  
:|:|:  
Db 68 YQMWF 73

RESULT 3  
US-08-969-683A-11  
Sequence 11, Application US/08969683A  
Patent No. 6136576  
GENERAL INFORMATION:  
APPLICANT: GENENCOR INTERNATIONAL, INC.

TITLE OF INVENTION: METHOD FOR THE RECOMBINANT  
TITLE OF INVENTION: PRODUCTION OF 1,3 PROPANEDIOL  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genencor International, Inc.  
STREET: 4 Cambridge Place  
STREET: 1870 South Winton road  
CITY: Rochester  
STATE: NY  
COUNTRY: U.S.A  
ZIP: 14618  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/969,683A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/20873  
FILING DATE: 13-NOV-1997  
APPLICATION NUMBER: 60/030,601  
FILING DATE: 13-NOV-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Glaister, Debra  
REGISTRATION NUMBER: 33,888  
REFERENCE/DOCKET NUMBER: GC 369-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-864-7620  
TELEFAX: 650-845-6504  
TELEX:  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: GPD1  
US-08-969-683A-11

Query Match 85.3%; Score 29; DB 4; Length 391;  
Best Local Similarity 50.0%; Pred. No. 8.2e+02;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 IQIWF 6  
:|:|:  
Db 68 YQMWF 73

RESULT 4  
US-09-297-928-7  
Sequence 7, Application US/09297928  
Patent No. 6358716  
GENERAL INFORMATION:  
APPLICANT: BUTTHUIS, BEN A.  
GATENBY, ANTHONY A.  
HAYNIE, SHARON L.  
HSU, AMY K.  
LAREAU, RICHARD D.  
TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF  
GLYCEROL BY RECOMBINANT  
ORGANISMS  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENENCOR INTERNATIONAL, INC.  
STREET: 4 CAMBRIDGE PLACE  
STREET: 1870 SOUTH WINTON ROAD  
CITY: ROCHESTER  
STATE: NEW YORK

COUNTRY: U.S.A.  
ZIP: 14618  
COMPUTER READABLE FORM:  
MEDIUM TYPE: DISKETTE, 3.5 INCH  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: MICROSOFT WINDOWS 95  
SOFTWARE: MICROSOFT WORD VERSION 7.0A  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/297,928  
FILING DATE: 11-May-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/03602  
FILING DATE: NOVEMBER 13, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: FLOYD, LINDA AXAMETHY  
REGISTRATION NUMBER: 33,692  
REFERENCE/DOCKET NUMBER: CR-9981-P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 302-892-8112  
TELEFAX: 302-773-0164  
TELEX: 6717325  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 391 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-297-928-7

Query Match 85.3%; Score 29; DB 4; Length 391;  
Best Local Similarity 50.0%; Pred. No. 8.2e+02;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IQIWF 6  
:|:|:  
Db 68 VQMWVF 73

RESULT 5  
US-08-919-624-1  
Sequence 1, Application US/08919624  
Patent No. 5994097  
GENERAL INFORMATION:  
APPLICANT: Lal, Preeti  
APPLICANT: Shah, Purvi  
APPLICANT: Guegler, Karl J.  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: NEW HUMAN G-PROTEIN COUPLED  
RECEPTOR  
TITLE OF INVENTION: RECEPTOR  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/919,624  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0377 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 433 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: CARDNOT01  
CLONE: 282414  
US-08-919-624-1

Query Match 85.3%; Score 29; DB 2; Length 433;  
Best Local Similarity 83.3%; Pred. No. 8.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IQIWF 6  
:|:|:  
Db 162 ISIWIF 167

RESULT 6  
US-09-025-151-18  
Sequence 18, Application US/09025151  
Patent No. 6187535  
GENERAL INFORMATION:  
APPLICANT: Legrain, Pierre  
APPLICANT: Fromont, Micheline  
APPLICANT: Rain, Jean-Christophe  
TITLE OF INVENTION: FAST AND EXHAUSTIVE METHOD FOR SELECTING A PREY  
TITLE OF INVENTION: POLYPEPTIDE INTERACTING WITH A BAIT POLYPEPTIDE OF  
TITLE OF INVENTION: INTEREST: APPLICATION TO THE CONSTRUCTION OF MAPS OF  
FILE REFERENCE: 03495-0164  
CURRENT APPLICATION NUMBER: US/09/025,151  
CURRENT FILING DATE: 1998-02-18  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 18  
LENGTH: 68  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-025-151-18

Query Match 82.4%; Score 28; DB 4; Length 68;  
Best Local Similarity 50.0%; Pred. No. 2.4e+02;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
:|:|:  
Db 7 IQVWLY 12

RESULT 7  
US-09-065-474-145  
Sequence 145, Application US/09065474  
Patent No. 6063599  
GENERAL INFORMATION:  
APPLICANT: Tang, Liang  
APPLICANT: Blehm, E. Scot  
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 171  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins

STATE: Colorado  
COUNTRY: USA  
ZIP: 80525  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: WordPerfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/065,474  
FILING DATE: 24-APR-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: HW-5-C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 145:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 287 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-065-474-145

Query Match 82.4%; Score 28; DB 3; Length 287;  
Best Local Similarity 66.7%; Pred. No. 8.9e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
Db 11 LRIMIF 16

RESULT 8  
US-09-557-034-145  
Sequence 145, Application US/09557034  
Patent No. 6365569  
GENERAL INFORMATION:  
APPLICANT: Tang, Liang  
Blehm, E. Scot  
TITLE OF INVENTION: DIOFILARIA AND BRUGIA ANKYRIN  
PROTEINS, NUCLEIC ACID MOLECULES, AND  
USES THEREOF  
NUMBER OF SEQUENCES: 171  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: WordPerfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/557,034  
FILING DATE: 21-APR-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/065,474  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: HW-5-C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272

TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 145:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 287 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 145:  
US-09-557-034-145

Query Match 82.4%; Score 28; DB 4; Length 287;  
Best Local Similarity 66.7%; Pred. No. 8.9e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
Db 11 LRIMIF 16

RESULT 9  
US-09-031-485-2  
Sequence 2, Application US/09031485  
Patent No. 582306  
GENERAL INFORMATION:  
APPLICANT: Tang, Liang  
Blehm, E. Scot  
TITLE OF INVENTION: DIOFILARIA AND BRUGIA ANKYRIN  
PROTEINS, NUCLEIC ACID MOLECULES, AND  
USES THEREOF  
NUMBER OF SEQUENCES: 85  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: WordPerfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/031,485  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/847,429  
FILING DATE: 24-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: HW-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 312 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-031-485-2

Query Match 82.4%; Score 28; DB 2; Length 312;  
Best Local Similarity 66.7%; Pred. No. 9.6e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
Db 159 LRIMIF 164

RESULT 10  
US-08-847-429A-2  
Sequence 2, Application US/08847429A  
Patent No. 5827692  
GENERAL INFORMATION:  
APPLICANT: Tang, Liang  
APPLICANT: Blehm, E. Scot  
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 85  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: WordPerfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/847,429A  
FILING DATE: 24-APR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: HW-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/493-7272  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 312 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-847-429A-2  
Query Match 82.4%; Score 28; DB 2; Length 312;  
Best Local Similarity 66.7%; Pred. No. 9.6e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
1 IQIWI 6  
:|||||  
Db 159 LRIWI 164

RESULT 11  
US-09-065-474-2  
Sequence 2, Application US/09065474  
Patent No. 6063599  
GENERAL INFORMATION:  
APPLICANT: Tang, Liang  
APPLICANT: Blehm, E. Scot  
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 171  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: WordPerfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/065,474  
FILING DATE: 24-APR-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: HW-5-C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/484-9505  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 312 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-065-474-2  
Query Match 82.4%; Score 28; DB 3; Length 312;  
Best Local Similarity 66.7%; Pred. No. 9.6e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
1 IQIWI 6  
:|||||  
Db 159 LRIWI 164

RESULT 12  
US-09-557-034-2  
Sequence 2, Application US/09557034  
Patent No. 6365569  
GENERAL INFORMATION:  
APPLICANT: Tang, Liang  
APPLICANT: Blehm, E. Scot  
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 171  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carol Talkington Verser, Ph.D.  
ADDRESSEE: Heska Corporation  
STREET: 1825 Sharp Point Drive  
CITY: Fort Collins  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80525  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: WordPerfect for Windows, Version 7.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/557,034  
FILING DATE: 21-APR-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/065,474  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Verser, Carol Talkington  
REGISTRATION NUMBER: 37,459  
REFERENCE/DOCKET NUMBER: HW-5-C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 970/484-9505  
TELEFAX: 970/484-9505  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 312 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2  
US-09-557-034-2

Query Match 82.4%; Score 28; DB 4; Length 312;  
Best Local Similarity 66.7%; Pred. NO. 9.6e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
DB 159 IQIWF 164

RESULT 13  
US-09-134-001C-5503  
Sequence 5503, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 5503  
LENGTH: 413  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-5503

Query Match 82.4%; Score 28; DB 4; Length 413;  
Best Local Similarity 80.0%; Pred. NO. 1.2e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 QIWF 6  
DB 110 QIWF 114

RESULT 14  
US-09-134-001C-3703  
Sequence 3703, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3703  
LENGTH: 496  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3703

Query Match 82.4%; Score 28; DB 4; Length 496;  
Best Local Similarity 50.0%; Pred. NO. 1.5e+03;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6

DB 157 MQIWF 162

RESULT 15  
US-09-134-001C-3275  
Sequence 3275, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 3275  
LENGTH: 591  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3275

Query Match 82.4%; Score 28; DB 4; Length 591;  
Best Local Similarity 66.7%; Pred. NO. 1.7e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 IQIWF 6  
DB 359 VNIWF 364

Search completed: January 3, 2003, 15:34:30  
Job time : 10 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: January 3, 2003, 15:32:11 ; Search time 4.77273 Seconds  
(without alignments)  
23.825 Million cell updates/sec

Title: US-09-543-188a-21

Perfect score: 34

Sequence: 1 IQIWI 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 117078 seqs, 18951520 residues 117078

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications-AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	85.3	179	US-09-789-561-92	Sequence 92, Appl
2	29	85.3	261	US-09-789-561-169	Sequence 169, App
3	28	82.4	82	US-09-864-761-47624	Sequence 47624, A
4	28	82.4	92	US-09-736-457-331	Sequence 331, App
5	28	82.4	92	US-09-902-941-331	Sequence 331, App
6	28	82.4	92	US-09-849-626-331	Sequence 331, App
7	28	82.4	95	US-09-925-300-1436	Sequence 1436, App
8	28	82.4	228	US-09-815-242-5172	Sequence 5172, App
9	28	82.4	233	US-09-741-669-319	Sequence 319, App
10	28	82.4	233	US-09-912-020-386	Sequence 386, App
11	28	82.4	233	US-09-815-242-10353	Sequence 10353, A
12	28	82.4	233	US-09-815-242-13969	Sequence 13969, A
13	28	82.4	234	US-09-815-242-11592	Sequence 11592, A
14	28	82.4	235	US-09-815-242-11103	Sequence 11103, A
15	28	82.4	293	US-09-738-626-3652	Sequence 3652, App
16	28	82.4	314	US-09-886-055-495	Sequence 495, App
17	27	79.4	39	US-09-764-877-1254	Sequence 1254, App
18	27	79.4	66	US-09-815-242-5869	Sequence 5869, App
19	27	79.4	87	US-09-796-692-1581	Sequence 1581, App

20	27	79.4	110	US-09-867-550-298	Sequence 298, App
21	27	79.4	264	US-09-925-301-905	Sequence 905, App
22	27	79.4	283	US-09-970-989-2	Sequence 2, Appl
23	27	79.4	366	US-09-925-301-1175	Sequence 1175, App
24	27	79.4	384	US-09-738-626-6166	Sequence 6166, App
25	27	79.4	411	US-09-901-252-16	Sequence 16, Appl
26	27	79.4	466	US-09-804-551B-14	Sequence 14, Appl
27	27	79.4	572	US-09-815-242-5626	Sequence 5626, App
28	27	79.4	573	US-09-815-242-12247	Sequence 12247, A
29	27	79.4	573	US-09-815-242-12921	Sequence 12921, A
30	27	79.4	608	US-09-924-336-8	Sequence 8, Appl
31	27	79.4	803	US-09-815-242-5597	Sequence 5597, App
32	27	79.4	805	US-09-815-242-12286	Sequence 12286, A
33	26	76.5	93	US-09-864-761-45734	Sequence 45734, A
34	26	76.5	111	US-09-925-299-790	Sequence 790, App
35	26	76.5	139	US-08-818-581B-11	Sequence 11, Appl
36	26	76.5	197	US-09-811-284-197	Sequence 197, App
37	26	76.5	242	US-08-818-581B-9	Sequence 9, Appl
38	26	76.5	245	US-08-818-581B-13	Sequence 13, Appl
39	26	76.5	245	US-09-970-989-4	Sequence 4, Appl
40	26	76.5	253	US-09-970-989-9	Sequence 9, Appl
41	26	76.5	253	US-09-895-913A-130	Sequence 130, App
42	26	76.5	263	US-09-860-670-115	Sequence 115, App
43	26	76.5	337	US-09-828-478-5	Sequence 5, Appl
44	26	76.5	337	US-09-866-230-8	Sequence 8, Appl
45	26	76.5	343	US-09-815-242-11887	Sequence 11887, A

## ALIGNMENTS

RESULT 1

US-09-789-561-92

Sequence 92, Application US/09789561

Patent No. US20020064818A1

GENERAL INFORMATION:

APPLICANT: NI et al.

TITLE OF INVENTION: 52 Human secreted proteins

FILE REFERENCE: P2043p1

CURRENT APPLICATION NUMBER: US/09/789, 561

CURRENT FILING DATE: 2001-02-22

PRIOR APPLICATION NUMBER: PCT/US00/24008

PRIOR FILING DATE: 2000-08-31

PRIOR APPLICATION NUMBER: 60/152, 317

PRIOR FILING DATE: 1999-09-03

PRIOR APPLICATION NUMBER: 60/152, 315

PRIOR FILING DATE: 1999-09-03

NUMBER OF SEQ ID NOS: 194

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 92

LENGTH: 179

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (143)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-789-561-92

Query Match

Best Local Similarity 85.3%; Score 29; DB 10; Length 179;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IQIWI 6

I I I I I

DB 86 INIWI 91

RESULT 2

US-09-789-561-169

Sequence 169, Application US/09789561

Patent No. US20020064818A1

GENERAL INFORMATION:

APPLICANT: Na et al.  
TITLE OF INVENTION: 52 Human secreted proteins  
FILE REFERENCE: P2043P1  
CURRENT APPLICATION NUMBER: US/09/789,561  
CURRENT FILING DATE: 2001-02-22  
PRIOR APPLICATION NUMBER: PCT/US00/24008  
PRIOR FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: 60/152,317  
PRIOR FILING DATE: 1999-09-03  
PRIOR APPLICATION NUMBER: 60/152,315  
PRIOR FILING DATE: 1999-09-03  
NUMBER OF SEQ ID NOS: 194  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 169  
LENGTH: 261  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (225)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
09-789-561-169

Query Match  
Best Local Similarity 85.3%; Score 29; DB 10; Length 261;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 IQIWF 6  
|:|:|:  
Db 168 IQIWF 173

RESULT 3  
US-09-864-761-47624  
Sequence 47624, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aemlica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 47624  
LENGTH: 82  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC011236.2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.5  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1  
OTHER INFORMATION: SWISSPROT HIT: P08578, EVALUATE 2.00e-16  
OTHER INFORMATION: EST\_HUMAN HIT: BF680487.1, EVALUATE 6.00e-31  
US-09-864-761-47624

Query Match  
Best Local Similarity 82.4%; Score 28; DB 10; Length 82;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
|:|:|:  
Db 54 IQVWLY 59

RESULT 4  
US-09-736-457-331  
Sequence 331, Application US/09736457  
Patent No. US20020168637A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Tonglong  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Lodes, Michael A.  
APPLICANT: Fanger, Gary  
APPLICANT: Vedvick, Tom  
APPLICANT: Carter, Darriick  
APPLICANT: Retter, Marc  
APPLICANT: Mannion, Jane  
APPLICANT: Fan, Liqun  
APPLICANT: Wang, Aijun  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.478C15  
CURRENT APPLICATION NUMBER: US/09/736,457  
CURRENT FILING DATE: 2000-12-13  
NUMBER OF SEQ ID NOS: 1864  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 331  
LENGTH: 92  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-736-457-331

Query Match  
Best Local Similarity 82.4%; Score 28; DB 9; Length 92;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IQIWF 6  
|:|:|:  
Db 31 IQVWLY 36

RESULT 5  
US-09-902-941-331  
; Sequence 331, Application US/09902941  
; Patent No. US20020172952A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Marnerakis, Margarita  
; APPLICANT: Carter, Darick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedavick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: McNabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C17  
; CURRENT APPLICATION NUMBER: US/09/902,941  
; CURRENT FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 331  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-902-941-331

Query Match 82.4%; Score 28; DB 9; Length 92;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 IOIWF 6  
||:|:|:  
DB 31 IOVWLX 36

RESULT 6  
US-09-849-626-331  
; Sequence 331, Application US/09849626  
; Publication No. US20020197669A1  
; GENERAL INFORMATION:  
; APPLICANT: Bangur, Chaitanya  
; APPLICANT: Fanger, Gary  
; APPLICANT: Wang, Aijun  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Switzer, Anne  
; APPLICANT: McNeill, Patricia  
; APPLICANT: Clapper, Jonathan  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C16  
; CURRENT APPLICATION NUMBER: US/09/849,626  
; CURRENT FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 1926  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 331  
; LENGTH: 92  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-849-626-331

Query Match 82.4%; Score 28; DB 9; Length 92;  
Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 IOIWF 6  
||:|:|:  
DB 31 IOVWLX 36

RESULT 7

US-09-925-300-1436  
; Sequence 1436, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben,  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1436  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-300-1436

Query Match 82.4%; Score 28; DB 10; Length 95;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 IOIWF 6  
||:|:|:  
DB 34 IOVWLX 39

RESULT 8  
US-09-815-242-5172  
; Sequence 5172, Application US/09815242  
; Patent No. US2002061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; TITLE OF INVENTION: Prokaryotes  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5172  
; LENGTH: 228  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-815-242-5172

Query Match 82.4%; Score 28; DB 10; Length 228;  
Best Local Similarity 50.0%; Pred. No. 3.1e+02;

Matches 3: Conservative 3: Mismatches 0: Indels 0: Gaps 0:

QY 1 IOIWF 6  
:::||||

Db 198 VKWIF 203

RESULT 9  
US-09-741-669-319  
; Sequence 319, Application US/09741669  
; Patent No. US2002022718A1  
; GENERAL INFORMATION:  
; APPLICANT: Forsyth, R. Allyn  
; APPLICANT: Ohlssen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; TITLE OF INVENTION: Genes identified as required for  
; FILE REFERENCE: ELITRA.009A  
; CURRENT FILING DATE: 2000-12-19  
; PRIOR FILING DATE: 2000-12-19  
; PRIOR APPLICATION NUMBER: US 60/173005  
; NUMBER OF SEQ ID NOS: 481  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 319  
; LENGTH: 233  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-741-669-319

Query Match 82.4%: Score 28; DB 10; Length 233;  
Best Local Similarity 50.0%: Pred. NO. 3.2e+02;  
Matches 3: Conservative 3: Mismatches 0: Indels 0: Gaps 0:

QY 1 IOIWF 6  
:::||||

Db 198 VKWIF 203

RESULT 10  
US-09-912-020-386  
; Sequence 386, Application US/09912020  
; Patent No. US20020045592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ohlssen, Judith  
; APPLICANT: Trawick, John  
; APPLICANT: Forsyth, R. Allyn  
; APPLICANT: Ficoelich, Jamie M.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN  
; FILE REFERENCE: ELITRA.001DVI  
; CURRENT FILING DATE: 2001-07-23  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: 09/492,709  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: 60/117,405  
; NUMBER OF SEQ ID NOS: 485  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 386  
; LENGTH: 233  
; TYPE: PRT  
; ORGANISM: E. coli  
US-09-912-020-386

Query Match 82.4%: Score 28; DB 10; Length 233;  
Best Local Similarity 50.0%: Pred. NO. 3.2e+02;  
Matches 3: Conservative 3: Mismatches 0: Indels 0: Gaps 0:

QY 1 IOIWF 6  
:::||||

Db 198 VKWIF 203

RESULT 11  
US-09-815-242-10353  
; Sequence 10353, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlssen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10353  
; LENGTH: 233  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-815-242-10353

Query Match 82.4%: Score 28; DB 10; Length 233;  
Best Local Similarity 50.0%: Pred. NO. 3.2e+02;  
Matches 3: Conservative 3: Mismatches 0: Indels 0: Gaps 0:

QY 1 IOIWF 6  
:::||||

Db 198 VKWIF 203

RESULT 12  
US-09-815-242-13969  
; Sequence 13969, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlssen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078

```
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 13969
;; LENGTH: 233
;; TYPE: PRF
;; ORGANISM: Salmonella typhi
09-815-242-13969
```

```
Query Match      82.4%; Score 28; DB 10; Length 233;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 IOIWI 6
      ::::|
Db      198 VKWIF 203
```

```
RESULT 13
US-09-815-242-11592
;; Sequence 11592, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; TITLE OF INVENTION: Prokaryotes
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 11592
;; LENGTH: 234
;; TYPE: PRF
;; ORGANISM: Helicobacter pylori
US-09-815-242-11592
```

```
Query Match      82.4%; Score 28; DB 10; Length 234;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 IOIWI 6
      ::::|
Db      198 VKWIF 203
```

```
RESULT 14
US-09-815-242-11103
;; Sequence 11103, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; TITLE OF INVENTION: Prokaryotes
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 11103
;; LENGTH: 235
;; TYPE: PRF
;; ORGANISM: Haemophilus influenzae
US-09-815-242-11103
```

```
Query Match      82.4%; Score 28; DB 10; Length 235;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 IOIWI 6
      ::::|
Db      198 VKWIF 203
```

```
RESULT 15
US-09-738-626-3652
;; Sequence 3652, Application US/09738626
;; Publication No. US20020197605A1
;; GENERAL INFORMATION:
;; APPLICANT: NAKAGAWA, SATOSHI
;; APPLICANT: MIZOGUCHI, HIROSHI
;; APPLICANT: ANDO, SEIKO
;; APPLICANT: HAYASHI, MIKIRO
;; APPLICANT: OCHIAI, KEIKO
;; APPLICANT: YOKOI, HARUHIKO
;; APPLICANT: TATEISHI, NAKO
;; APPLICANT: SENOH, AKIHIRO
;; APPLICANT: IKEDA, MASATO
;; APPLICANT: OZAKI, AKIO
;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;; FILE REFERENCE: 249-125
;; CURRENT APPLICATION NUMBER: US/09/738,626
```

; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; PRIOR FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 7059  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 3652  
 ; LENGTH: 293  
 ; TYPE: PRT  
 ; ORGANISM: Corynebacterium glutamicum  
 US-09-738-626-3652

Query Match 82.4%; Score 28; DB 9; Length 293;  
 Best Local Similarity 50.0%; Pred. No. 3.9e+02;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

1 IQIWF 6  
 :|:|:|  
 89 LQIWL 94

Search completed: January 3, 2003, 15:52:06  
 Job time : 4.77273 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:21:45 : Search time 26.5909 Seconds  
(without alignments)  
30.067 Million cell updates/sec

Title: US-09-543-188a-27  
Perfect score: 38  
Sequence: 1 WLWIA 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	6	23	AAU11850
2	38	100.0	371	23	AB847793
3	37	97.4	230	22	AA689817
4	37	97.4	230	22	AA876811
5	35	92.1	70	22	AAU47680
6	35	92.1	241	23	ABP27890
7	34	89.5	49	23	AB68706
8	34	89.5	80	22	AA66002
9	34	89.5	162	22	AAU32492
10	34	89.5	235	22	AAU57493

11	34	89.5	288	22	AAU42440	Protonbacterium
12	34	89.5	420	22	AAU35667	Haemophilus influe
13	34	89.5	676	20	AAV35418	Chlamydia pneumoni
14	33	86.8	38	21	AA851380	Human secreted pep
15	33	86.8	78	22	AA002505	Human polypeptide
16	33	86.8	95	23	ABP07995	Human ORFX protein
17	33	86.8	132	22	ABG10849	Novel human diagno
18	33	86.8	145	22	AAU53601	Protonbacterium
19	33	86.8	162	21	AA655841	Arabidopsis thalia
20	33	86.8	171	21	AA655840	Arabidopsis thalia
21	33	86.8	171	21	AA650774	Arabidopsis thalia
22	33	86.8	185	21	AA655839	Arabidopsis thalia
23	33	86.8	185	21	AA650773	Arabidopsis thalia
24	33	86.8	296	22	AA698330	Escherichia coli p
25	33	86.8	329	22	AA696276	Putative P. abyss
26	33	86.8	362	15	AA688695	G-protein coupled
27	33	86.8	362	17	AAW02657	G-protein coupled
28	33	86.8	396	22	ABG24925	Novel human diagno
29	33	86.8	472	22	ABG24925	Novel human diagno
30	33	86.8	497	20	ABG10850	Protein which is s
31	33	86.8	504	22	ABG10850	Novel human diagno
32	33	86.8	735	22	AAU36496	Pseudomonas aerugi
33	33	86.8	735	22	AAU36496	Human 5' EST seque
34	32	84.2	32	21	AAV11779	Hepatitis GB virus
35	32	84.2	75	21	ABP09351	Human ORFX protein
36	32	84.2	84	23	ABP02429	Human polypeptide
37	32	84.2	89	22	AA600323	Human secreted pro
38	32	84.2	95	21	AA600323	Human ORFX protein
39	32	84.2	105	23	ABP01702	Human normal bladd
40	32	84.2	115	20	AAV60396	Sequence of the VL
41	32	84.2	128	15	AA654053	p64-h2 protein pro
42	32	84.2	139	13	AA629009	Human immune/haema
43	32	84.2	149	22	AA690245	Arabidopsis thalia
44	32	84.2	153	21	AA615863	Arabidopsis thalia
45	32	84.2	153	21	AA621898	Arabidopsis thalia

#### ALIGNMENTS

RESULT 1	AAU11850	standard; peptide: 6 AA.
ID	AAU11850	
XX	AC	AAU11850:
XX	DT	26-MAR-2002 (first entry)
XX	XX	Peptide ligand for Prion protein, PrP, #25.
DE	XX	
XX	KW	Prion protein; PrP; ligand; octapeptide motif; scrapie;
KW	KW	prion-associated disease; Creutzfeldt-Jakob disease;
KW	KW	Gerstmann-Strausler-Scheinker disease; fatal familial insomnia;
KW	KW	feline spongiform encephalopathy; bovine spongiform encephalopathy;
KW	KW	transmissible mink encephalopathy; exotic ungulate encephalopathy;
XX	XX	chronic wasting disease.
OS	XX	Synthetic.
PN	XX	MO20017687-A2.
PD	XX	18-OCT-2001.
XX	PF	05-APR-2001; 2001WO-US11150.
XX	XX	
PR	XX	05-APR-2000; 2000US-0543188.
XX	XX	(VITE-) VI TECHNOLOGIES INC.
PA	XX	
XX	XX	Streptococcus poly
PI	XX	Human prostate spe
XX	XX	Human immune/haema
XX	XX	Novel human secret
DR	XX	Protonbacterium
XX	XX	WPI; 2002-061944/08.

PT New ligands for prion proteins, useful for detection or removal or  
PT prions and for treating prion-associated diseases, recognize a specific  
PT octapeptide motif -  
XX  
XX  
PS Claim 16; Page 34; 47pp; English.  
XX  
CC The invention relates to a ligand of less than 6 kD that binds to a  
CC polypeptide containing the sequence GLYTYRGLYGLINPROHISGLYGLY (A) or an  
CC analogue that is the retro-inverso isomer of (A). The sequence A is  
CC an octapeptide motif from the prion protein (Prp). The ligands are  
CC identified by binding assays with the peptide (A) or peptides containing  
CC (A). The ligands are used for detecting prion proteins (or prions) in  
CC biological or environmental samples, e.g. for diagnosis, also for  
CC removing them from samples, for treating or retarding development of  
CC prion-associated diseases, especially Creutzfeldt-Jakob diseases (in  
CC iatrogenic, new variant, familial or sporadic forms), but also  
CC Gerstmann-Strausler-Scheinker disease, fatal familial insomnia,  
CC scrapie, bovine or feline spongiform encephalopathy, transmissible mink  
CC or exotic ungulate encephalopathy, or chronic wasting disease. The  
CC present sequence is a ligand of the invention.

## Sequence 6 AA:

Query Match 100.0%; Score 38; DB 23; Length 6;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWVIA 6  
| | | | | |  
Db 1 WLWVIA 6

## RESULT 2

ABBA47793 standard; protein; 371 AA.

AC ABBA47793;

DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #497.

KW Antlacterial; gene therapy; vaccine; biosynthesis; biodegradation;

KM vitamin B12; bacterial infection; disease.

OS Listeria monocytogenes.

XX MO200177335-A2.

PN 18-OCT-2001.

PF 11-APR-2001; 2001MO-FR01118.

PR 11-APR-2000; 2000FR-0004629.

PA (INSP ) INST PASTEUR.

PI Buchrieser C, Frangoul L, Couve E, Rusniok C, Eshti H, Dehoux P,  
PI Dussanget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;  
PI Daniels J, Goebel W, Kieft J, Kunh M, Ng E, Vazquez-Boland JA;  
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;  
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;  
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;  
PI Madueno E, De Pablos B, Wehlend J, Kaerst U, Entian K, Hauf J;  
PI Rose M, Voss H;

PI MPI; 2002-010914/01.

PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment  
PT and prevention of Listeria and related bacterial infections, and  
PT related polypeptides -

PS Claim 6; SEQ ID No 498; 192pp; French.

XX  
CC The present invention relates to the genome sequence of Listeria  
CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of  
CC it are useful for selecting probes and primers for detecting genes in L.  
CC monocytogenes and other organisms, and for studying genetic  
CC polymorphisms and other genomes. The present sequence is a protein  
CC encoded by the genome sequence of the present invention. Proteins  
CC expressed from the genome sequence are useful for raising specific  
CC antibodies, identification of L. monocytogenes and related organisms, and  
CC for biosynthesis and biodegradation, especially biosynthesis of vitamin  
CC B12. The genome sequence and proteins encoded by it are also useful for  
CC selecting compounds that regulate gene expression and cell replication  
CC and modulate L. monocytogenes-related diseases. In addition, the genome  
CC sequence and proteins encoded by it are useful in pharmaceutical and  
CC vaccines compositions for the treatment or prevention of infections by L.  
CC monocytogenes and related organisms.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

## SQ Sequence 371 AA:

Query Match 100.0%; Score 38; DB 23; Length 371;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWVIA 6  
| | | | | |  
Db 313 WLWVIA 318

## RESULT 3

AAG89817 standard; protein; 230 AA.

AC AAG89817;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 3571.

KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KM organic acid synthesis.

OS Corynebacterium glutamicum.

XX EP1108790-A2.

PN 20-JUN-2001.

PF 18-DEC-2000; 2000EP-0127688.

PR 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.

PR 03-AUG-2000; 2000JP-0280988.

PA (KYOW ) KYOWA HAKKO KOGYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

PI MPI; 2001-376931/40.

DR N-PSDB; AAH65036.

PT Novel polynucleotides derived from Coryneform bacteria, for identifying  
PT mutation point of a gene, measuring expression of a gene, analysing  
PT expression profile or pattern of a gene and identifying homologous gene

PS Claim 17; SEQ ID NO: 3571; 246pp + Sequence Listing; English.

CC The present invention provides a number of nucleotide and protein  
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These



CC are useful for identifying the mutation point of a gene derived from a  
 CC mutant of corynebacterium bacterium, measuring expression amount and  
 CC analysing the expression profile or expression pattern of a gene derived  
 CC from corynebacterium bacterium, and identifying a homologue of a gene derived  
 CC from corynebacterium bacterium. Corynebacterium bacteria are useful for producing  
 CC amino acids, nucleic acids, vitamins, saccharides and organic acids,  
 CC particularly L-lysine. The present sequence is a protein described  
 CC in the exemplification of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC European Patent Office.

CC Sequence 230 AA;

Query Match 97.4%; Score 37; DB 22; Length 230;  
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 WLWIA 6  
 170 WLWIA 175

RESULT 4  
 AAB76811  
 ID AAB76811 standard; Protein; 230 AA.

AC AAB76811;

DT 11-APR-2001 (first entry)

DE Corynebacterium glutamicum MCT protein SEQ ID NO:604.

XX Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;  
 KW membrane construction and membrane transport protein; petrioleum spill;  
 KW hydrocarbon degradation; gram positive aerobic bacterium; marker;  
 KW identification; microorganism; fine chemical production; transformation;  
 KW genome mapping; genetic engineering.

XX OS Corynebacterium glutamicum.

PN WO200100805-A2.

PD 04-JAN-2001.

PF 23-JUN-2000; 2000WO-IB00926.

XX 25-JUN-1999; 99US-0141031.  
 PR 08-JUL-1999; 99DE-1031454.  
 PR 08-JUL-1999; 99DE-1031478.  
 PR 08-JUL-1999; 99DE-1031563.  
 PR 09-JUL-1999; 99DE-1032122.  
 PR 09-JUL-1999; 99DE-1032124.  
 PR 09-JUL-1999; 99DE-1032125.  
 PR 09-JUL-1999; 99DE-1032128.  
 PR 09-JUL-1999; 99DE-1032180.  
 PR 09-JUL-1999; 99DE-1032182.  
 PR 09-JUL-1999; 99DE-1032190.  
 PR 09-JUL-1999; 99DE-1032191.  
 PR 09-JUL-1999; 99DE-1032209.  
 PR 09-JUL-1999; 99DE-1032212.  
 PR 09-JUL-1999; 99DE-1032227.  
 PR 09-JUL-1999; 99DE-1032228.  
 PR 09-JUL-1999; 99DE-1032229.  
 PR 09-JUL-1999; 99DE-1032230.  
 PR 14-JUL-1999; 99DE-1032927.  
 PR 14-JUL-1999; 99DE-1033005.  
 PR 27-AUG-1999; 99DE-1040764.  
 PR 27-AUG-1999; 99DE-1040765.  
 PR 27-AUG-1999; 99DE-1040766.  
 PR 27-AUG-1999; 99DE-1040830.  
 PR 27-AUG-1999; 99DE-1040831.

PR 27-AUG-1999; 99DE-1040832.  
 PR 27-AUG-1999; 99DE-1040833.  
 PR 31-AUG-1999; 99DE-1041378.  
 PR 31-AUG-1999; 99DE-1041379.  
 PR 31-AUG-1999; 99DE-1041395.  
 PR 03-SEP-1999; 99DE-1042077.  
 PR 03-SEP-1999; 99DE-1042078.  
 PR 03-SEP-1999; 99DE-1042079.  
 PR 03-SEP-1999; 99DE-1042088.

XX (BADI ) BASF AG.

XX Pompejus M, Kroege B, Schroeder H, Zelder O, Haberhauer G;

XX WPI: 2001-071486/08.  
 DR N-PSDB: AAF68044.

PT Corynebacterium glutamicum nucleic acids encoding membrane construction  
 PT and membrane transport proteins or their portions, useful for typing or  
 PT identifying C. glutamicum or related bacteria, and as markers for  
 PT transformation

PS Claim 20; Page 1012; 1119pp; English.

CC AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane  
 CC construction and membrane transport (MCT) proteins given in AAB76510 to  
 CC AAB76847. The MCT nucleic acids and proteins are useful in the  
 CC identification of microorganisms which can be used to produce fine  
 CC chemicals, for modulating fine chemical production in C. glutamicum or  
 CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or  
 CC identification of C. glutamicum or related bacteria, as reference points  
 CC for mapping C. glutamicum genome, and as markers for transformation.  
 CC AAF68082 and AAF68082 represent sequencing primers which are used in an  
 CC example from the present invention.

XX Sequence 230 AA;

Query Match 97.4%; Score 37; DB 22; Length 230;  
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWIA 6  
 170 WLWIA 175

RESULT 5  
 AAU47680  
 ID AAU47680 standard; Protein; 70 AA.

AC AAU47680;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #8576.

KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.  
 PR 02-JUN-2000; 2000US-208841P.  
 PR 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIXA CORP.  
XX  
XX Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A:  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D:  
XX  
XX WPI: 2001-616774/71.  
DR N-PSDB: AAS59539.  
XX

PT Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris -

PS Example 1: SEQ ID No 8875; 1069pp; English.

XX  
XX Sequences AAU9105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 70 AA:

Query Match 92.1%; Score 35; DB 22; Length 70;  
Best Local Similarity 83.3%; Pred. No. 78;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWMTA 6  
1:|||||  
Db 26 WVVMTA 31

RESULT 6

ABP27890 ABP27890 standard; Protein: 241 AA.

ABP27890;

XX 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 4956.

XX  
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
XX

OS Streptococcus agalactiae.

PN WO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

XX 24-NOV-2000; 2000GB-0028727.

XX 07-MAR-2001; 2001GB-0005640.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.  
XX  
XX Telford J, Masiyanni V, Margarit Ros YI, Grandi G, Fraser C:  
PI Tetteilin H;  
XX  
XX WPI: 2002-352536/38.  
DR N-PSDB: ABN68521.  
XX

PT New Streptococcus protein for the treatment or prevention of infection  
PT or disease caused by Streptococcus bacteria, such as meningitis, and  
PT for detecting a compound that binds to the protein -

PS Claim 1: Page 3656; 4525pp; English.

XX  
XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5403 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins.  
XX

SQ Sequence 241 AA:

Query Match 92.1%; Score 35; DB 23; Length 241;  
Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWMTA 6  
1:|||||  
Db 168 WLMWTA 173

RESULT 7

ABG68706 ABG68706 standard; Protein: 49 AA.

ABG68706;

XX 07-OCT-2002 (first entry)

XX Human prostate specific protein DEX0293\_75.

XX  
XX Human; prostate specific nucleic acid; PSNA; prostate cancer; PSP;  
KW prostate specific protein; cytosolic; non-cancerous prostate disease;  
KW gene therapy; cancer; immunostimulant; vaccine.  
XX

OS Homo sapiens.

PN WO200255735-A2.

XX 18-JUL-2002.

XX 27-NOV-2001; 2001WO-US44363.

XX 27-NOV-2000; 2000US-253176P.

XX (DIAD-) DIADEXUS INC.

XX Salceda S, Macina RA, Recipon H, Cafferey R, Ali S, Sun Y;  
PI Liu C, Chen S;

XX WPI: 2002-557831/59.

DR N-PSDB; ABK97583.  
XX New prostate specific genes, useful for treating or diagnosing cancer,  
PT or useful as vaccines for treating cancer, particularly prostate  
PT cancer, in a patient  
XX  
PS Claim 11; Page 192; 212pp; English.  
XX  
CC The invention relates to a new isolated prostate-specific nucleic acid  
CC (PSNA) molecule comprising the cDNA sequences appearing as ABK97574-  
CC ABK97642 which encode prostate specific proteins appearing as  
CC ABG68701-ABG68746, or a sequence hybridizing to a PSNA or which has 60%  
CC sequence homology with a PSNA. Also included are a method of determining  
CC the presence of a PSNA in a sample, a vector comprising the PSNA,  
CC a host cell comprising the vector, producing the polypeptide encoded by  
CC the PSNA, a method of determining the presence of a PSP in a sample,  
CC diagnosing and monitoring the presence and metastases of prostate cancer  
CC in a patient, a kit for detecting a risk of cancer or presence of cancer  
CC in a patient (the kit comprising a means for determining the presence of  
CC the PSNA or PSP in a sample of a patient) and a vaccine comprising the  
CC polypeptide or the nucleic acid encoding the polypeptide. The PSNA, PSP  
CC and anti-PSP antibody are useful for diagnosing and treating cancer in a  
CC patient (e.g. by gene therapy). The nucleic acid molecule and polypeptide  
CC are also useful as vaccines for treating cancer, particularly prostate  
CC cancer and non-cancerous prostate diseases. The present sequence is  
CC a PSP of the invention.  
XX  
SQ Sequence 49 AA;  
  
Query Match 89.5%; Score 34; DB 23; Length 49;  
Best Local Similarity 66.7%; Pred. No. 77;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 WLWVTA 6  
|:|:|:  
8 WVWVVA 13  
Db  
  
RESULT 8  
AAM86002  
ID AAM86002 standard; Protein; 80 AA.  
XX  
AC AAM86002;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen SEQ ID NO:13595.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytostatic; gene therapy; vaccine; metastasis.  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
PD  
XX 09-AUG-2001.  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0188874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225265.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0233397.  
PR 14-SEP-2000; 2000US-0233398.  
PR 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-0234200.  
PR 14-SEP-2000; 2000US-0234201.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 13-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249399.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
  
(HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI: 2001-483426/52.  
XX N-PSDB: AAK58783.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
XX Claim 11: SEQ ID NO 13595; 3071pp + Sequence listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting

CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169  
CC represent sequences used in the exemplification of the present invention.  
XX  
XX  
SQ Sequence 80 AA:  
  
Query Match 89.5%; Score 34; DB 22; Length 80;  
Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 WLVVIA 6  
||:|:|  
Db 61 WLVVIA 66  
  
RESULT 9  
AAU32492  
ID AAU32492 standard; Protein; 162 AA.  
XX  
XX AC AAU32492;  
XX  
XX DT 18-DEC-2001 (first entry)  
XX  
XX DE Novel human secreted protein #2983.  
XX  
XX KW Human; vaccination; gene therapy; nutritional supplement;  
KW stem cell proliferation; hematopoiesis; nerve tissue regeneration;  
KW immune suppression; immune stimulation; anti-inflammatory; leukemia.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200179449-A2.  
XX  
XX PD 25-OCT-2001.  
XX  
XX PF 16-APR-2001; 2001WO-US08656.  
XX  
XX PR 18-APR-2000; 2000US-0552929.  
XX PR 26-JAN-2001; 2001US-0770160.  
XX  
XX PA (HYSE-) HYSEQ INC.  
XX  
XX PI Tang YT, Liu C, Dzmanac RT;  
XX WPI: 2001-611725/70.  
XX  
XX Nucleic acids encoding a range of human polypeptides, useful in genetic  
PT vaccination, testing and therapy -  
XX  
XX Claim 20: Page 623; 765pp; English.  
XX  
XX The invention relates to novel human secreted polypeptides. The  
CC polypeptides and antibodies to the polypeptides are useful for  
CC determining the presence of or predisposition to a disease associated  
CC with altered levels of polypeptide. The polypeptides are also useful for  
CC identifying agents (agonists and antagonists) that bind to them. Cells  
CC expressing the proteins are useful for identifying a therapeutic agent  
CC for use in treatment of a pathology related to aberrant expression or  
CC physiological interactions of the polypeptide. Vectors comprising  
CC the nucleic acids encoding the polypeptides and cells genetically  
CC engineered to express them are also useful for producing the proteins.  
CC The proteins are useful in genetic vaccination, testing and  
CC therapy, and can be used as nutritional supplements. They may be used to  
CC increase stem cell proliferation; to regulate haematopoiesis; and in  
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;  
CC immune suppression and/or stimulation; as anti-inflammatory agents; and  
CC in treatment of leukemias. AAU29510-AAU33304 represent the amino acid  
XX sequences of novel human secreted proteins of the invention.

S0 Sequence 162 AA;  
 Query Match 89.5%; Score 34; DB 22; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
 QY 1 WLWWT 5  
 11111  
 Db 16 WLWWT 20  
 RESULT 10  
 AA057493  
 ID AA057493 standard; Protein; 235 AA.  
 XX AA057493;  
 XX  
 DT 13-FEB-2002 (first entry)  
 Propionibacterium acnes immunogenic protein #18389.  
 KW SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 OS Propionibacterium acnes.  
 XX  
 XX MO200181581-h2.  
 XX  
 XX 01-NOV-2001.  
 PD  
 PE 20-APR-2001; 2001MO-US12865.  
 XX  
 PR 21-APR-2000; 2000US-199047P.  
 PR 02-JUN-2000; 2000US-208841P.  
 PR 07-JUL-2000; 2000US-216747P.  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX  
 DR WPI; 2001-616774/71.  
 DR N-PSDB; AAS59583.  
 XX  
 PM  
 PP  
 PS  
 PS Example 1: SEQ ID NO 16688; 1069pp; English.  
 XX  
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hyperostosis and osteomyelitis) uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 at ftp.wipo.int/pub/published\_pct\_sequences.

Seq	Sequence	235 AA:
QY	1 WLWVIA 6 1 1111	89.5%; Score 34; DB 22; Length 235; Best Local Similarity 83.3%; Pred. No. 3.9e+02; Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0
Db	182 WFWVIA 187	
RESULT 11		
AAU42440		
ID	AAU42440 standard; Protein; 288 AA.	
XX		
AC	AAU42440;	
XX		
DT	27-FEB-2002 (first entry)	
XX		
DE	Propionibacterium acnes immunogenic protein #3336.	
XX		
KM	SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;	
KM	uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;	
KM	inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;	
KM	dermatological; osteopathic; neuroprotectant.	
XX		
OS	Propionibacterium acnes.	
XX		
PN	WO200181581-A2.	
XX		
XX	01-NOV-2001.	
PD	20-APR-2001; 2001WO-US12865.	
XX		
PF	21-APR-2000; 2000US-199047P.	
PR	02-JUN-2000; 2000US-208841P.	
PR	07-JUL-2000; 2000US-216747P.	
PA	(CORI-) CORIXA CORP.	
XX		
PI	Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;	
PI	L'maisonneuve J, Zhang Y, Jen S, Carter D;	
XX		
DR	WPI: 2001-616774/71.	
XX	N-PSDB; AAS59517.	
PT	Propionibacterium acnes polypeptides and nucleic acids useful for	
PT	vaccinating against and diagnosing infections, especially useful for	
PT	treating acne vulgaris -	
XX		
PS	Example 1; SEQ ID NO 3635; 1066pp; English.	
XX		
CC	Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic	
CC	polypeptides. The proteins and their associated DNA sequences are used in	
CC	the treatment, prevention and diagnosis of medical conditions caused by	
CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,	
CC	pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.	
CC	P. acnes is also involved in infections of bone, joints and the central	
CC	nervous system, however it is particularly involved in the inflammatory	
CC	lesions associated with acne vulgaris. A method for detecting the	
CC	presence or absence of P. acnes in a patient comprises contacting a	
CC	sample with a binding agent that binds to the proteins of the invention	
CC	and determining the amount of bound protein in the sample. The	
CC	polypeptides may be used as antigens in the production of antibodies	
CC	specific for P. acnes proteins. These antibodies can be used to	
CC	downregulate expression and activity of P. acnes polypeptides and	
CC	therefore treat P. acnes infections. The antibodies may also be used as	
CC	diagnostic agents for determining P. acnes presence, for example, by	
CC	enzyme linked immunosorbent assay (ELISA).	
CC	Note: The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX		

SQ Sequence 288 AA;  
Query Match 89.5%; Score 34; DB 22; Length 288;  
Best Local Similarity 100.0%; Pred. No. 4.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WLWMI 5  
IIIII  
Db 126 WLWMI 130

RESULT 12  
AAU35667  
ID AAU35667 standard; Protein: 420 AA.  
XX  
AC AAU35667;  
XX  
DT 14-FEB-2002 (first entry)  
XX  
DE Haemophilus influenzae cellular proliferation protein #308.  
XX  
NM Antisense: prokaryotic cellular proliferation protein;  
XX antibiotic; antibacterial; drug design.  
XX  
OS Haemophilus influenzae.  
XX  
PN WO200170955-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 21-MAR-2001; 2001WO-US09180.  
XX  
PR 21-MAR-2000; 2000US-191078P.  
XX 23-MAY-2000; 2000US-206848P.  
PR 26-MAY-2000; 2000US-207272P.  
XX 23-OCT-2000; 2000US-242578P.  
PR 27-NOV-2000; 2000US-253625P.  
XX 22-DEC-2000; 2000US-257931P.  
PR 16-FEB-2001; 2001US-269308P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
PI Yamamoto RT, Xu HH;  
XX  
DR WPI: 2001-611495/70.  
DR N-PSDB: AAS53526.  
XX  
XX New polynucleotides for the identification and development of  
antibiotics, comprise sequences of antisense nucleic acids -  
Example 3; Seq ID No 11260; 511pp; English.

PS The invention relates to antisense inhibitors of genes essential to  
XX prokaryotic cellular proliferation, their use in identifying the  
XX genes, their use in the discovery of novel antibiotics, the essential  
XX genes themselves and the encoded proteins. The prokaryotes used are  
XX *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
XX *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
XX invention is also useful for the identification of potential new targets  
XX for antibiotic development. The antisense nucleic acids can also be used  
XX to identify proteins used in proliferation, to express these proteins,  
XX and to obtain antibodies capable of binding to the expressed proteins.  
XX The proteins can be used to screen compounds in rational drug discovery  
XX programmes. The antisense nucleic acid sequence is also useful to screen  
XX for homologous nucleic acids which are required for cell proliferation in  
XX a wide variety of organisms. The present sequence represents an  
XX essential prokaryotic cellular proliferation protein.  
XX Note: The sequence data for this patent did not form part  
XX of the printed specification, but was obtained in electronic  
XX format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pcl\_sequences.

SQ Sequence 420 AA;  
Query Match 89.5%; Score 34; DB 22; Length 420;  
Best Local Similarity 100.0%; Pred. No. 7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WLWMI 5  
IIIII  
Db 10 WLWMI 14

RESULT 13  
AA53418  
ID AA53418 standard; Protein: 676 AA.  
XX  
AC AA53418;  
XX  
DT 13-SEP-1999 (first entry)  
XX  
DE Chlamydia pneumoniae cellular envelope protein.  
XX  
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
KW vaccine; neutralising epitope.  
XX  
OS Chlamydia pneumoniae.  
XX  
PN WO927105-A2.  
XX  
PD 03-JUN-1999.  
XX  
PF 20-NOV-1998; 98WO-IB01890.  
XX  
PR 04-NOV-1998; 98US-0107078.  
XX 21-NOV-1997; 97FR-0014673.  
XX  
PA (GEST) GENSET.  
XX  
PI Griffais R;  
XX  
DR WPI: 1999-357842/30.  
XX  
XX Genome sequence of Chlamydia pneumoniae  
XX  
XX  
XX  
PS Page 1199-1201; Disclosure; 1912pp; English.

XX AAY34584-Y35879 represent the proteins encoded by all the open reading  
XX frames in the complete genome (see AAY34584-Y35879) of Chlamydia pneumoniae.  
XX C. pneumoniae causes respiratory disease such as pneumonia and  
XX bronchitis and is thought to be a contributing factor in heart  
XX disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
XX nodosum or pharyngitis. The polypeptides encoded by the open reading  
XX frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in  
XX immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
XX nucleotide sequences can also be used as immunogenic compositions,  
XX especially where the vector directs the expression of a neutralising  
XX epitope of C. pneumoniae.

SQ Sequence 676 AA;  
Query Match 89.5%; Score 34; DB 20; Length 676;  
Best Local Similarity 100.0%; Pred. No. 1.1e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WLWMI 5  
IIIII  
Db 481 WLWMI 485

RESULT 14  
AAB51390  
ID AAB51390 standard; Peptide: 38 AA.  
XX

AC	AA051390;
XX	
DT	16-FEB-2001 (first entry)
DE	
XX	Human secreted peptide #5 encoded by cDNA #11.
XX	
KW	Cytostatic; immunosuppressive; neutrotropic; neuroprotective; antiviral;
KW	antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW	vulnerable; anticoagulant; antibacterial; antifungal; antiparasitic;
KW	candidant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW	neurological disease; infection; human; secreted protein.
OS	Homo sapiens.
XX	
PN	WO20058495-A1.
XX	
PD	05-OCT-2000.
XX	
PT	23-MAR-2000; 2000WO-US07661.
PR	26-MAR-1999; 99US-0126504.
XX	
PA	07-JAN-2000; 2000US-0174847.
XX	
PB	(HUMA-) HUMAN GENOME SCI INC.
PI	Rosen CA, Ruben SM, Komatsoulis G;
XX	
DR	WPI: 2000-611720/58.
DR	N-PSDB; AAC93320.
PT	New nucleic acid molecules encoding 45 human secreted proteins for
PT	diagnosing, preventing, treating or ameliorating medical conditions and
XX	used as food additives or preservatives -
PS	Claim 11; Pages 356-357; 410pp: English.
XX	
CC	Sequences AA051380-B51A23 represent the amino acid sequences of 45
CC	human secreted proteins encoded by the genes AAC93310-C93354. The genes
CC	and proteins are useful for preventing, ameliorating or treating medical
CC	conditions, e.g. by protein or gene therapy. The genes are isolated from
CC	a range of human tissues disclosed in the specification. The nucleic
CC	acids, proteins, antibodies and (ant)agonists are useful in the
CC	diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC	ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC	marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC	(b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC	haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC	Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC	colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC	(d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC	epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC	and parasitic infections.
SO	Sequence 38 AA:
OY	Query Match 86.8%; Score 33; DB 21; Length 38;
Db	Best Local Similarity 80.0%; Pred. NO. 84;
	Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
	1 WLWMT 5
	16 WLWVW 20
RESULT 15	
ID	AA002505
XX	AA002505 standard; Protein: 78 AA.
XX	
DT	06-NOV-2001 (first entry)
XX	
DE	Human polypeptide SEQ ID NO 16397

```

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; Leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
XX
XX 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI: 2001-514838/56.
XX
XX N-PSDB; AA182436.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
XX diagnosing and treating e.g. leukaemia, inflammation and immune
XX disorders -
XX
XX Claim 20; SEQ ID NO 16397; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AA179941-AA19341) and
XX the encoded proteins (AA000010-AA013910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 78 AA:
XX
XX Query Match 86.8%; Score 33; DB 22; Length 78;
XX Best Local Similarity 80.0%; Pred. No. 1.8e+02;
XX Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 WLWMI 5
XX |||||
XX Db 31 WLWV 35

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:00 : Search time 9.55455 Seconds  
(without alignments)  
57.944 Million cell updates/sec

Title: US-09-543-188A-27  
Perfect score: 38  
Sequence: 1 WLWVIA 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Reached: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	38	100.0	371	2 AE1190	probable membrane
2	38	100.0	371	2 AE1548	probable membrane
3	37	97.4	719	2 B95325	conserved hypothet
4	36	94.7	651	2 AP2493	hypothetical prote
5	35	92.1	468	2 AD0521	probable symporter
6	35	92.1	536	2 T20736	hypothetical prote
7	35	92.1	565	2 S73854	hypothetical prote
8	35	92.1	585	2 T19814	hypothetical prote
9	34	89.5	285	2 E70473	phosphate transpor
10	34	89.5	318	2 G72310	conserved hypothet
11	34	89.5	384	2 T44870	hypothetical prote
12	34	89.5	420	2 AE4132	thiol-disulfide int
13	34	89.5	587	2 C85044	hypothetical prote
14	34	89.5	703	2 H72034	hypothetical prote
15	34	89.5	703	2 H86588	thiol-disulfide int
16	34	89.5	714	2 G81503	thiol-disulfide int
17	34	89.5	1394	2 S66876	ATP-dependent tran
18	33	86.8	164	2 T06913	H+-transporting tw
19	33	86.8	172	2 E70057	hypothetical prote
20	33	86.8	179	2 E84126	hypothetical prote
21	33	86.8	188	2 D97260	probable membrane
22	33	86.8	196	2 C84919	hypothetical prote
23	33	86.8	232	2 A95303	probable Aqp2 aqu
24	33	86.8	295	2 AC0499	probable phosphate
25	33	86.8	296	1 BVECP1	phosphate transpor
26	33	86.8	296	2 F91211	hypothetical prote
27	33	86.8	296	2 F86057	hypothetical prote
28	33	86.8	296	2 AD0956	phosphate transpor
29	33	86.8	329	2 E75073	hypothetical prote

30	33	86.8	338	2 E83688	hypothetical prote
31	33	86.8	355	2 C71092	hypothetical prote
32	33	86.8	391	2 C87615	acyltransferase, p
33	33	86.8	436	2 T36706	hypothetical prote
34	33	86.8	527	2 A82431	sensor protein Unp
35	33	86.8	594	2 T31824	hypothetical prote
36	33	86.8	528	2 F82895	hypothetical prote
37	33	86.8	638	2 T24676	hypothetical prote
38	33	86.8	692	2 C81653	probable thiol-dis
39	33	86.8	692	2 H71494	probable thiol-dis
40	33	86.8	699	2 H87275	thio-disulfide int
41	33	86.8	735	2 A83006	hypothetical prote
42	33	86.8	815	2 E83294	probable acyl-CoA
43	33	86.8	1411	2 S48442	Pdrl1 protein - ye
44	32	84.2	40	2 S08038	hypothetical prote
45	32	84.2	170	2 E75017	crRNA intron endonu

## ALIGNMENTS

RESULT 1  
AE1190  
probable membrane protein lmo0925 [imported] - Listeria monocytogenes (strain EGD-e)  
C:Species: Listeria monocytogenes  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AE1190  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.;  
Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AE1190  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-371 <GLA>  
A:Cross-References: GB:NC\_003210; PIDN:CAC99003.1; PID:g16410327; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo0925

Query Match 100.0%; Score 38; DB 2; Length 371;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
Db 313 WLWVIA 318

RESULT 2  
AE1548  
probable membrane protein lin0925 [imported] - Listeria innocua (strain Clp11262)  
C:Species: Listeria innocua  
C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AE1548  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.;  
Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AE1548  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-371 <GLA>  
A:Cross-References: GB:AL592022; PIDN:CAC96157.1; PID:g16413373; GSPDB:GN00178  
A:Experimental source: strain Clp11262

C:Genetics:  
A:Gene: 11n0925

Query Match 100.0%; Score 38; DB 2; Length 371;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
|||||  
DB 313 WLWVIA 318

RESULT 3  
B95325

conserved hypothetical protein Sma0937 [imported] - *Sinorhizobium meliloti* (strain 1021)  
C:Species: *Sinorhizobium meliloti*  
C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: B95325  
R:Barrett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows  
R:Barrett, M.J.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.  
J: Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A:Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti*  
A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: B95325  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-719 <KUR>  
A:Cross-references: GB:AE006469; PIDN:AK65164.1; PID:g14523607; GSPDB:GN00165  
A:Experimental source: strain 1021, megaplasmid pSymA  
R:Gallibert, F.; Flman, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
P.; Chalm, P.; Cowley, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; LeLaure,  
hepault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: Sma0937  
A:Genome: plasmid

Query Match 97.4%; Score 37; DB 2; Length 719;  
Best Local Similarity 83.3%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
|||||  
DB 333 WLWVIA 338

RESULT 4  
AF2493

hypothetical protein alr7126 [imported] - *Nostoc* sp. (strain PCC 7120) plasmid pCC7120a1  
C:Species: *Nostoc* sp.  
A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C:Accession: AF2493  
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi,  
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium *Ana*  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AF2493  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-651 <KUR>  
A:Cross-references: GB:BA000020; PIDN:BA078210.1; PID:g17135664; GSPDB:GN00180  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr7126  
A:Genome: plasmid

Query Match 94.7%; Score 36; DB 2; Length 651;  
Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
|||||  
DB 369 WLWVIA 374

RESULT 5  
AD0521

probable symporter STY0171 [imported] - *Salmonella enterica* subsp. *enterica* serovar T  
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh  
A:Note: this species has also been called *Salmonella typhi*  
C>Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AD0521  
R:Parikh, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Fair  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* se  
A:Reference number: AB0502; PMID:11677608  
A:Accession: AD0521  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-468 <PAR>  
A:Cross-references: GB:AL13382; PIDN:CAD01308.1; PID:g16501436; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY0171  
C:Superfamily: melbiose carrier protein

Query Match 92.1%; Score 35; DB 2; Length 468;  
Best Local Similarity 83.3%; Pred. No. 1.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
|||||  
DB 82 WLWVIA 87

RESULT 6  
T20736

hypothetical protein F11A10.5 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Mar-2000  
C:Accession: T20736  
R:Smyle, R.  
submitted to the EMBL Data Library, December 1995  
A:Reference number: T20736  
A:Accession: T20736  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-536 <WIL>  
A:Cross-references: EMBL:568297; PIDN:CAA92595.1; GSPDB:GN00022; CESP:F11A10.5  
A:Experimental source: clone F11A10  
C:Genetics:  
A:Gene: CESP:F11A10.5  
A:Map position: 4  
A:Introns: 36/1; 74/2; 100/1; 157/1; 237/1; 269/3; 340/2  
C:Superfamily: *Caenorhabditis elegans* hypothetical protein F11A10.5

Query Match 92.1%; Score 35; DB 2; Length 536;  
Best Local Similarity 83.3%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
|||||  
DB 9 WLWVIA 14

RESULT 7  
S73854

hypothetical protein F10orf565 - Mycoplasma pneumoniae (strain ATCC 29342)  
C:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C:Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999  
C:Accession: S73854  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A:Reference number: S73327; MUID:97105885; PMID:8948633  
A:Accession: S73854  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-565 <NTM>  
A:Cross-references: EMBL:AE000052; GB:U00089; NID:q1674223; PIDN:AB96176.1; PID:q167422  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C:Genetics:  
A:Genetic code: SGC3

Query Match 92.1%; Score 35; DB 2; Length 565;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
|||:|  
Db 137 WLWVIA 142

RESULT 8  
T19814  
hypothetical protein C38C6.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000  
C:Accession: T19814  
R:Baynes, C.  
submitted to the EMBL Data Library, March 1997  
A:Reference number: Z19182  
A:Accession: T19814  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-585 <WIL>  
A:Cross-references: EMBL:Z93375; PIDN:CAB07562.1; GSPDB:GN00020; CESP:C38C6.2  
C:Genetics:  
A:Gene: CESP:C38C6.2  
A:Map position: 2  
A:Insertions: 17/2; 97/2; 218/1; 299/2; 337/1; 372/3; 517/3; 547/3  
A:Superfamily: alpha-glucosidase; alpha-amylase core homology

Query Match 92.1%; Score 35; DB 2; Length 585;  
Best Local Similarity 66.7%; Pred. No. 2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
|||:|  
Db 105 WLWVIA 110

RESULT 9  
E70473  
phosphate transport system permease PstA - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 26-Aug-1999  
C:Accession: E70473  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: E70473  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-285 <AGF>  
A:Cross-references: GB:AE000768; NID:g2984249; PIDN:AAC07781.1; PID:g2984254; GB:AE00065

A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: pstA  
C:Superfamily: phoM protein

Query Match 89.5%; Score 34; DB 2; Length 285;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVI 5  
|||||  
Db 35 WLWVI 39

RESULT 10  
G72310  
conserved hypothetical protein - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: G72310  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Ginn, M.L.; Dodson, R.J.; Haft, D.H.; Hic  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,  
C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: G72310  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-318 <ARN>  
A:Cross-references: GB:AE001760; GB:AE000512; NID:g4981510; PIDN:AA036065.1; PID:g498  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM0986

Query Match 89.5%; Score 34; DB 2; Length 318;  
Best Local Similarity 83.3%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
|||||  
Db 125 WLWVIA 130

RESULT 11  
T44870  
acylttransferase homolog [imported] - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 18-Feb-2000  
C:Accession: T44870  
R:Parkhill, J.; Barrell, B.G.; Rastandream, M.A.  
submitted to the EMBL Data Library, April 1998  
A:Reference number: Z22863  
A:Accession: T44870  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-384 <PAR>  
A:Cross-references: EMBL:AL022486; PIDN:CAA18555.1  
A:Experimental source: cosmid B1883  
C:Genetics:  
A:Note: MICB1883 06  
C:Superfamily: 4'''-O-acylttransferase

Query Match 89.5%; Score 34; DB 2; Length 384;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
|||:|  
Db 120 WLWVIA 125

RESULT 12

A64132  
amphotropic murine retrovirus receptor homolog - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 08-Oct-1999  
C:Accession: A64132  
R:Flitschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.; Glodex, A.; Kelley, J.M.; Weidman, J.  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.  
Science 269, 436-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630; PMID:7542800  
A:Accession: A64132  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-420 <TIGR>  
A:Cross-references: GB:U32834; GB:I42023; NID:g1574444; PIDN:AAC23248.1; PID:g1574446; T

Query Match 89.5%; Score 34; DB 2; Length 420;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWMI 5  
|||||  
Db 10 WLWMI 14

RESULT 13  
C85044  
hypothetical protein AT4g03490 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: C85044  
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488; PMID:10617198  
A:Accession: C85044  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-587 <STO>  
A:Cross-references: GB:NC\_001268; NID:g7270672; PIDN:CAB77834.1; GSPDB:GN00140  
C:Genetics:  
A:Gene: AT4g03490  
A:Map position: 4

Query Match 89.5%; Score 34; DB 2; Length 587;  
Best Local Similarity 66.7%; Pred. No. 2.8e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWMI 6  
|||||  
Db 312 WLWMI 317

RESULT 14  
H72034  
thiol-disulfide interchange protein dsbd [imported] - Chlamydomonas reinhardtii (strain C  
C:Species: Chlamydomonas reinhardtii  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 28-Jul-2000  
C:Accession: H72034  
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydomonas reinhardtii and C. trachomatis.  
A:Reference number: A72000; MUID:99206606; PMID:10192388  
A:Accession: H72034  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-703 <ARR>  
A:Cross-references: GB:AE001660; GB:AE001363; NID:g4377095; PIDN:AAD18924.1; PID:g4377095  
C:Genetics:  
A:Gene: dsbd

Query Match 89.5%; Score 34; DB 2; Length 703;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWMI 5  
|||||  
Db 508 WLWMI 512

RESULT 15  
H86588  
thio-disulfide interchange protein [imported] - Chlamydomonas reinhardtii (strain J138  
C:Species: Chlamydomonas reinhardtii  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
C:Accession: H86588  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of Chlamydomonas reinhardtii J138.  
A:Reference number: A86491; MUID:20330349; PMID:10871362  
A:Accession: H86588  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-703 <STO>  
A:Cross-references: GB:BA000008; NID:g8979159; PIDN:BA098994.1; GSPDB:GN00142  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: dsbd

Query Match 89.5%; Score 34; DB 2; Length 703;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWMI 5  
|||||  
Db 508 WLWMI 512

Search completed: January 3, 2003, 15:33:20  
Job time : 10.9545 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: January 3, 2003, 15:22:10 ; Search time 5.18182 Seconds

(without alignments)  
48.025 Million cell updates/sec

Title: US-09-543-188a-27

Perfect score: 38  
Sequence: 1 WLWIA 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	92.1	565	1 YD08_MYCPN	P75472 mycoplasma
2	34	89.5	318	1 Y966_THEMA	Q9X081 thermotoga
3	34	89.5	420	1 YG04_HAEIN	P45268 haemophilus
4	33	86.8	164	1 ATPX_CVAPA	P48085 cyanophora
5	33	86.8	295	1 PSTA_YERPE	P58655 yersinia pe
6	33	86.8	296	1 PSTA_YERPE	P07654 escherichia
7	33	86.8	355	1 YAU0_PYROH	Q58728 pyrococcus
8	33	86.8	506	1 NU2C_NEPOL	Q9L107 neptroselm
9	33	86.8	1410	1 PDRB_YEAST	P40550 saccharomyc
10	32	84.2	40	1 YPS1_PLEBO	P15733 plectonema
11	32	84.2	115	1 MERT_SHEPU	Q54462 shewanella
12	32	84.2	215	1 ERD2_PETHY	Q9Z4N2 petunia hyb
13	32	84.2	228	1 AQP2_BRUAB	Q91A79 brucella ab
14	32	84.2	228	1 AQP2_BRUAB	Q91A72 brucella me
15	32	84.2	232	1 Y836_AQUAE	P06708 aquifex aeo
16	32	84.2	269	1 TOX1_BORPE	P04877 bordetella
17	32	84.2	329	1 EXB8_PSEPU	Q05605 pseudomonas
18	32	84.2	389	1 GSP1_AERHY	P45782 aeromonas h
19	32	84.2	407	1 GSP1_AERHY	P45782 aeromonas h
20	32	84.2	438	1 Y281_HAEIN	P44610 haemophilus
21	32	84.2	510	1 MVIN_HAEIN	P44658 haemophilus
22	32	84.2	538	1 THIP_HAEIN	P44885 haemophilus
23	32	84.2	745	1 RNR_BUCAT	P57628 buchnera ap
24	32	84.2	837	1 AT54_HUMAN	Q24738 hom sapien
25	32	84.2	893	1 BOSS_DROVI	Q24738 drosophila
26	32	84.2	2283	1 DPOE_MOUSE	Q9W4F7 mus muscula
27	32	84.2	2286	1 DPOE_MOUSE	Q9W4F7 mus muscula
28	31	81.6	78	1 VG9_SPLYR	P15900 splioplasm
29	31	81.6	154	1 GIB2_NIPBR	P51535 nipostrom
30	31	81.6	243	1 YKJA_BACSU	P49653 bacillus su
31	31	81.6	341	1 YF71_PYROH	Q05927 pyrococcus
32	31	81.6	374	1 TR15_FUSSP	P13513 fusarium sp
33	31	81.6	375	1 TR15_GIBZE	Q00909 gibberella

## ALIGNMENTS

RESULT 1	ID	YD08_MYCPN	STANDARD:	PRT:	565 AA.
AC	P75472	008088:			
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DE	Hypothetical protein MPN308 (F10_orf565).				
GN	MPN308 OR MP528.				
OS	Mycoplasma pneumoniae.				
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.				
OX	NCBI_TaxID=2104;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 29342 / M129;				
RX	MEDLINE=97105885; PubMed=8948633;				
RA	Himmelfreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,				
RA	Herrmann R.;				
RT	"Complete sequence analysis of the genome of the bacterium Mycoplasma				
RT	pneumoniae.";				
RL	Nucleic Acids Res. 24:4420-4449(1996).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=ATCC 29342 / M129;				
RX	MEDLINE=97252497; PubMed=9098066;				
RA	Krause D.C., Proft T., Hedreyda C.T., Hilbert H., Plagens H.,				
RA	Herrmann R.;				
RT	"Transposon mutagenesis reinforces the correlation between Mycoplasma				
RT	pneumoniae cytoskeletal protein HMW2 and cytolysis.";				
RL	J. Bacteriol. 179:2668-2677(1997).				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).				
CC	-1- SIMILARITY: IN THE N-TERMINAL SECTION: TO M.PNEUMONIAE MPN095.				
CC	-1- SIMILARITY: IN THE C-TERMINAL SECTION: TO M.PNEUMONIAE MPN095.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	-----				
DR	EMBL: AE000052; AA096176.1;				
DR	EMBL: U59896; AA052525.1;				
DR	InterPro: IPR002293; AA/re1_pmeasel.				
KW	Hypothetical protein; Transmembrane; Complete proteome.				
FT	TRANSMEM 8				
FT	TRANSMEM 43				
FT	TRANSMEM 63				
FT	TRANSMEM 95				
FT	TRANSMEM 137				
FT	TRANSMEM 157				
FT	TRANSMEM 167				
FT	TRANSMEM 167				
FT	TRANSMEM 227				
FT	TRANSMEM 247				
FT	TRANSMEM 268				
FT	TRANSMEM 314				
FT	TRANSMEM 334				
FT	TRANSMEM 367				
FT	TRANSMEM 424				
FT	TRANSMEM 444				

000835 fusarium po  
P58957 drosophila  
P26769 gibberella  
O59947 stachybotry  
O13489 myrothecium  
P50537 schizosacch  
O60603 home sapien  
O95m53 macaca fasc  
P22815 drosophila  
P12722 avian infec  
P11223 avian infec  
P12650 avian infec

FT TRANSMEM 460 480 POTENTIAL.  
FT TRANSMEM 482 502 POTENTIAL.  
FT TRANSMEM 516 536 POTENTIAL.  
SQ SEQUENCE 565 AA: 62560 MW: ACC745543D2184E2 CRC64:

Query Match 92.1%; Score 35; DB 1; Length 565;  
Best Local Similarity 66.7%; Pred. No. 65;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWVIA 6  
DB 137 WLWVIA 142

RESULT 2  
Y986\_THEMA  
ID Y986\_THEMA STANDARD: PRT: 318 AA.  
AC Q9X081;  
DT 16-OCT-2001 (Rel. 40; Created)  
DT 16-OCT-2001 (Rel. 40; Last sequence update)  
DT 16-OCT-2001 (Rel. 40; Last annotation update)  
Hypothetical protein TM0986.  
GN TM0986.

OS Thermotoga maritima.  
OC Bacteria; Thermotogae; Thermotogales;  
OC Thermotogaceae; Thermotoga.  
OX NCBI\_TaxID=2336;

RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
Salberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
RT Evidence for lateral gene transfer between Archaea and Bacteria from  
genome sequence of Thermotoga maritima."  
RL Nature 399:323-329(1999).

CC -1- SIMILARITY: STRONG, TO E.COLI YFAT AND P.AERUGINOSA PA4450.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: AE001760; AAD36065.1; -  
DR TIGR: TM0986; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 318 AA: 36705 MW: 7041AB7510531380 CRC64:

Query Match 89.5%; Score 34; DB 1; Length 318;  
Best Local Similarity 83.3%; Pred. No. 57;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLWVIA 6  
DB 125 WLWVIA 130

RESULT 3  
YG04\_HAEIN  
ID YG04\_HAEIN STANDARD: PRT: 420 AA.  
AC P45268;  
DT 01-NOV-1995 (Rel. 32; Created)  
DT 01-NOV-1995 (Rel. 32; Last sequence update)  
DT 15-JUN-2002 (Rel. 41; Last annotation update)  
DE Putative phosphate permease HII604.  
GN HII604.

OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;

RA Keriavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick E.F.,  
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
Weidman J.F., Phillips C.A., Spiggs T., Hedblom E., Cotton M.D.,  
Utterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,  
Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
Rd."  
RL Science 269:496-512(1995).

CC -1- FUNCTION: POTENTIAL TRANSPORTER FOR PHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
(Potential).

CC -1- SIMILARITY: BELONGS TO THE PHO-4 FAMILY OF TRANSPORTERS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U32834; AAC23248.1; -  
DR TIGR: HII604; -  
DR InterPro: IPR001204; Phos\_transp.  
DR Pfam: PF01384; PHO4; 1.  
KW Hypothetical protein; Phosphate transport; Transport; Transmembrane;  
KW Inner membrane; Complete proteome.

FT TRANSMEM 8 28 POTENTIAL.  
FT TRANSMEM 49 69 POTENTIAL.  
FT TRANSMEM 88 108 POTENTIAL.  
FT TRANSMEM 112 132 POTENTIAL.  
FT TRANSMEM 145 165 POTENTIAL.  
FT TRANSMEM 185 205 POTENTIAL.  
FT TRANSMEM 216 236 POTENTIAL.  
FT TRANSMEM 250 270 POTENTIAL.  
FT TRANSMEM 300 320 POTENTIAL.  
FT TRANSMEM 343 363 POTENTIAL.  
FT TRANSMEM 370 390 POTENTIAL.  
FT TRANSMEM 393 413 POTENTIAL.  
SQ SEQUENCE 420 AA: 44026 MW: 89651207F01D3810 CRC64:

Query Match 89.5%; Score 34; DB 1; Length 420;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWVI 5  
DB 10 WLWVI 14

RESULT 4  
ATPX\_CYAPA  
ID APTX\_CYAPA STANDARD: PRT: 164 AA.  
AC P48085;  
DT 01-FEB-1996 (Rel. 33; Created)  
DT 01-FEB-1996 (Rel. 33; Last sequence update)  
DT 15-JUN-2002 (Rel. 41; Last annotation update)  
DE ATP synthase B' chain (EC 3.6.3.14) (Subunit II).  
GN APTG.  
OS Cyanophora paradoxa.  
OG Cyanelle.

```

OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.
ON NCBI_TaxID=2762;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LB555 / Pringsheim;
RA Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohert H.J.,
RT Bryant D.A.;
RL "Nucleotide sequence of the cyanelle DNA from Cyanophora paradoxa.";
RN Plant Mol. Biol. Rep. 13:327-332(1995).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=LB555 / Pringsheim;
RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
RA Farley J.Y., Schlachter W.M., Chung S., Neumann-Spallart C.,
RA Steiner J.M., Jakowitsch J., Bohert H.J., Bryant D.A.;
RT "The complete sequence of the cyanelle genome of Cyanophora paradoxa:
RT the genetic complexity of a primitive plastid.";
RL (in) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
RL Schwemmler W. (eds.);
RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
(1997).
CC
CC -1- FUNCTION: THE B'-SUBUNIT IS A DIVERGED AND DUPLICATED FORM OF
CC B FOUND IN PLANTS AND PHOTOSYNTHETIC BACTERIA.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC H(+)(Out).
CC -1- SUBUNIT: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE ATPASE COMPLEX.
CC -1- SUBCELLULAR LOCATION: CYANELLE THYLAKOID MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE ATPASE B CHAIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U30821; AAA81256.1; -.
CC InterPro: IPR002146; ATPsynth_B/B'sub.
CC Pfam: PF00430; ATP-synt_B; 1.
CC Hydrogen ion transport; Transmembrane; CF(0); Cyanelle.
CC SEQUENCE 164 AA; 18568 MW; 49ACCE15AF010D7C CRC64;
SQ

```

Query Match Best Local Similarity 86.8%; Score 33; DB 1; Length 164;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

Db 1 WLWVIA 6
   ||:|:|
   4 WIIWLA 9

```

RESULT 5  
PSTA\_YERPE STANDARD; PRT; 295 AA.  
AC P58655;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE Phosphate transport system permease protein pstA.  
GN PSTA OR YPO4115.  
OS Yerania pestis.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Yerania.  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CO-92 / Biovar Orientalis;  
RA MEDLINE=21470413; PubMed=11586360;  
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
RA Prentice M.B., Seabright M., James K.D., Churcher C., Nungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,

```

RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yerania pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
CC
CC -1- FUNCTION: PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC FOR PHOSPHATE; PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE
CC SUBSTRATE ACROSS THE MEMBRANE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -1- SIMILARITY: BELONGS TO THE BINDING-PROTEIN-DEPENDENT TRANSPORT
CC SYSTEM PERMEASE FAMILY. CYSTM SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: A0414160; CAC93564.1; -.
CC InterPro: IPR000515; BPD_transp.
CC TIGRFAMS: TIGR00974; 3a0107s02c; 1.
CC DR PROSITE: PS00402; BPD_TRANS_P. INN MEMBR. 1.
CC KM Transport; Phosphate transport; Transmembrane; Inner membrane;
CC Complete proteome.
CC FT TRANSMEM 29
CC FT TRANSMEM 88 110 POTENTIAL.
CC FT TRANSMEM 123 145 POTENTIAL.
CC FT TRANSMEM 150 169 POTENTIAL.
CC FT TRANSMEM 201 223 POTENTIAL.
CC FT TRANSMEM 268 290 POTENTIAL.
CC SQ SEQUENCE 295 AA; 32331 MW; EEE8B558CB2721EA CRC64;

```

Query Match Best Local Similarity 86.8%; Score 33; DB 1; Length 295;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 WLWVI 5
   ||:|:|
Db 44 WLWVI 48

```

RESULT 6  
PSTA\_ECOLI STANDARD; PRT; 296 AA.  
AC P07654;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1988 (Rel. 07, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Phosphate transport system permease protein pstA.  
GN PSTA OR PHOT OR B3726.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=85293094; PubMed=2993631;  
RA Anemura M., Makino K., Shingawa H., Kobayashi A., Nakata A.;  
RT "Nucleotide sequence of the genes involved in phosphate transport and  
RT regulation of the phosphate regulon in Escherichia coli.";  
RL J. Mol. Biol. 184:241-250(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=85104705; PubMed=3881386;  
RA Surin B.P., Rosenberg H., Cox G.B.;  
RT "Phosphate-specific transport system of Escherichia coli: nucleotide  
RT sequence and gene-polypeptide relationships.";  
RL J. Bacteriol. 161:189-198(1985).  
RN [3]

RP SEQUENCE FROM N.A.  
RX STRAIN-K12 / MG1655;  
RC MEDLINE-93315143; PubMed-766882;  
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;  
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli  
genome: organizational symmetry around the origin of replication.";   
RL Genomics 16:551-561(1993).  
RN [4]  
RP MUTAGENESIS.  
RX MEDLINE-93077562; PubMed-1447208;  
RA Webb D.C., Rosenberg H., Cox G.B.;  
RT "Mutational analysis of the Escherichia coli phosphate-specific  
transport system, a member of the traffic ATPase (or ABC) family of  
membrane transporters. A role for proline residues in transmembrane  
helices.";  
RL J. Biol. Chem. 267:24661-24668(1992).  
CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
FOR PHOSPHATE. PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE  
SUBSTRATE ACROSS THE MEMBRANE.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
CC -1- SIMILARITY: BELONGS TO THE BINDING-PROTEIN-DEPENDENT TRANSPORT  
SYSTEM PERMEASE FAMILY. CYSTM SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: X02723; CAA26508.1; -;  
DR EMBL: K01992; AAA24380.1; -;  
DR EMBL: L10328; AAA62077.1; -;  
DR EMBL: AE000449; AAC76749.1; -;  
DR PIR: B23311; BVCEPT.  
DR EcoGene: EG10782; psta.  
DR InterPro: IPR000515; BPD\_transp.  
DR Pfam: PF00528; BPD\_transp. 1.  
DR TrIRPAMS: TIGR00974; 3a0107s02c. 1.  
DR PROSITE: PS00402; BPD\_TRANS\_P\_NN\_MEMBR. 1.  
KW Transport; Phosphate transport; Transmembrane; Inner membrane;  
KM Complete proteome.  
FT DOMAIN 1 28 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 29 50 PERIPLASMIC.  
FT DOMAIN 51 82 PERIPLASMIC (POTENTIAL).  
FT TRANSMEM 83 102 POTENTIAL.  
FT DOMAIN 103 126 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 127 146 POTENTIAL.  
FT DOMAIN 147 150 PERIPLASMIC (POTENTIAL).  
FT TRANSMEM 151 169 POTENTIAL.  
FT DOMAIN 170 204 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 205 223 POTENTIAL.  
FT DOMAIN 224 266 PERIPLASMIC (POTENTIAL).  
FT TRANSMEM 267 286 POTENTIAL.  
FT DOMAIN 289 296 CYTOPLASMIC (POTENTIAL).  
SQ SEQUENCE 296 AA; 32321 MW; BFEADDEC0B9851DC7 CRC64;  
  
Query Match 86.8%; Score 33; DB 1; Length 296;  
Best Local Similarity 80.0%; Pred. No. 77;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 1 WLWMI 5  
Db 44 WLWMI 48  
  
RESULT 7  
ID YA00\_PYRHO STANDARD; PRT; 355 AA.  
AC 058728;  
DT 16-OCT-2001 (Rel. 40, Created)  
RX 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein PH1000.  
GN PH1000.  
OS Pyrococcus horikoshii.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=53953;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN-OT3;  
RC MEDLINE-98344137; PubMed-9679194;  
RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,  
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
RA Masuchi Y., Shizuya H., Kikuchi H.,  
RT "Complete sequence and gene organization of the genome of a hyper-  
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: BELONGS TO THE UPF0118 (PERM) FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: AP000004; BAA30097.1; -;  
DR InterPro: IPR002549; UPF0118.  
DR Pfam: PF01594; UPF0118; 1.  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 34 54 POTENTIAL.  
FT TRANSMEM 55 75 POTENTIAL.  
FT TRANSMEM 84 104 POTENTIAL.  
FT TRANSMEM 158 178 POTENTIAL.  
FT TRANSMEM 212 232 POTENTIAL.  
FT TRANSMEM 240 260 POTENTIAL.  
FT TRANSMEM 274 294 POTENTIAL.  
FT TRANSMEM 310 330 POTENTIAL.  
SQ SEQUENCE 355 AA; 40179 MW; D0929BBFEA847F39 CRC64;  
  
Query Match 86.8%; Score 33; DB 1; Length 355;  
Best Local Similarity 50.0%; Pred. No. 90;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 1 WLWMI 6  
Db 258 WLWMI 263  
  
RESULT 8  
ID NU2C\_NEPOL STANDARD; PRT; 506 AA.  
AC 09T107;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE NADH-plastoquinone oxidoreductase chain 2, chloroplast (EC 1.6.5.3).  
GN NDH.  
OS Nephroselmis olivacea.  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Chlorophyta; Prasinophyceae;  
OC Chlorodendrales; Chlorodendraceae; Nephroselmis.  
OX NCBI\_TaxID=31312;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-NIES-484;  
MEDLINE-9398694; PubMed-10468594;  
RA Turnell M., Otis C., Lemieux C.;



"The complete chloroplast DNA sequence of the green alga *Nephroselmis olivacea*: insights into the architecture of ancestral chloroplast genomes.";  
 Proc. Natl. Acad. Sci. U.S.A. 96:10248-10253(1999).  
 CC -1- CARBOLYTIC ACTIVITY: NADH + plastoquinone = NAD(+) + plastoquinol.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast.  
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 2 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF137379; AAD54809.1; -;  
 DR InterPro: IPR001750; Oxidored\_q1.  
 DR Pfam: PF00361; Oxidored\_q1; 1.  
 CC Oxidoreductase; NAD; Plastoquinone; Chloroplast; Transmembrane.  
 FT TRANSMEM 15 34 POTENTIAL.  
 FT TRANSMEM 41 63 POTENTIAL.  
 FT TRANSMEM 85 104 POTENTIAL.  
 FT TRANSMEM 117 139 POTENTIAL.  
 FT TRANSMEM 172 194 POTENTIAL.  
 FT TRANSMEM 215 237 POTENTIAL.  
 FT TRANSMEM 247 269 POTENTIAL.  
 FT TRANSMEM 282 304 POTENTIAL.  
 FT TRANSMEM 338 360 POTENTIAL.  
 FT TRANSMEM 380 402 POTENTIAL.  
 FT TRANSMEM 417 436 POTENTIAL.  
 FT TRANSMEM 469 491 POTENTIAL.  
 SQ SEQUENCE 506 AA; 54338 MW; 073F8705BFFA22A7 CRC64;  
 Query Match 86.8%; Score 33; DB 1; Length 506;  
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 WLVWIA 6  
 Db 42 WLVWIA 47  
 RESULT 9  
 PDBR\_YEAST STANDARD; PRT; 1410 AA.  
 ID PDBR\_YEAST 003092;  
 AC P40550; Q03092;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE ATP-dependent permease PDR11.  
 GN PDR11 OR YIL013C.  
 GN Saccharomyces cerevisiae (Baker's yeast).  
 OS Saccharomycetes; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / AB972;  
 RA Barrett B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,  
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,  
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,  
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,  
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,  
 RA Walsh S.V., Whitehead S.;  
 RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RP IDENTIFICATION AND SEQUENCE OF 1-10.  
 RC MEDLINE=95355421; PubMed=7629127;  
 RA Decotlines A., Lambert L., Catly P., Degand H., Epping E.A.,  
 RA Moye-Rowley W.S., Balzi E., Goffeau A.;  
 RT Identification and characterization of SNO2, a new multidrug ATP binding cassette transporter of the yeast plasma membrane.";

J. Biol. Chem. 270:18150-18157(1995).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. PDR5 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: 238113; CAA86236.1; -;  
 DR EMBL: 246881; CAA86980.1; -;  
 DR SGD: S0001275; PDR11.  
 DR InterPro: IPR003439; ABC\_transport.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD000006; ABC\_transporter; 2.  
 DR ProSITE: PS00211; ABC\_TRANSPORTER; 1.  
 KW ATP-binding; Transmembrane; Glycoprotein; Transport.  
 FT INIT MET 0  
 FT DOMAIN 1 387 POTENTIAL.  
 FT TRANSMEM 388 408 POTENTIAL.  
 FT TRANSMEM 418 438 POTENTIAL.  
 FT TRANSMEM 471 491 POTENTIAL.  
 FT TRANSMEM 494 514 POTENTIAL.  
 FT TRANSMEM 524 544 POTENTIAL.  
 FT TRANSMEM 636 656 POTENTIAL.  
 FT DOMAIN 657 1089 POTENTIAL.  
 FT TRANSMEM 1090 1110 POTENTIAL.  
 FT TRANSMEM 1117 1137 POTENTIAL.  
 FT TRANSMEM 1175 1195 POTENTIAL.  
 FT TRANSMEM 1204 1224 POTENTIAL.  
 FT TRANSMEM 1230 1250 POTENTIAL.  
 FT TRANSMEM 1335 1375 POTENTIAL.  
 FT TRANSMEM 1376 1410 POTENTIAL.  
 FT DOMAIN 1378 788 POTENTIAL.  
 FT NP\_BIND 781 788 ATP (POTENTIAL).  
 FT DOMAIN 707 712 POLY-SER.  
 FT DOMAIN 1045 1048 POLY-LEU.  
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1288 1288 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1323 1323 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1345 1345 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1410 AA; 160405 MW; 93C9399A5CD114C3 CRC64;  
 Query Match 86.8%; Score 33; DB 1; Length 1410;  
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 WLVWIA 6  
 Db 552 WLVWIA 557  
 RESULT 10  
 YPSL\_PLEBO STANDARD; PRT; 40 AA.  
 ID YPSL\_PLEBO  
 AC P15733;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-APR-1990 (Rel. 14, Last annotation update)  
 DE Hypothetical 4.6 kDa protein (ORF1).  
 OS Plectonema boryanum.  
 OS Plectonema small.  
 OS Plasmid small.  
 OC Bacteria; Cyanobacteria; Oscillatoriales; Plectonema.  
 OX NCBI\_TaxID=1184;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 27894 / UTEX 581;  
 RA Wickrema A., Barnum S.R., Jaworski J.G.;  
 RL Submitted (JAN-1989) to the EMBL/GenBank/DBJ databases.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X51320; -; NOT\_ANNOTATED\_CDs.  
CC DR PIR: S08038; S08038.  
CC KM Hypothetical protein; Plasmid.  
CC SQ SEQUENCE 40 AA: 4592 MW: 83D6BFBDB296D7D37 CRC64;  
  
Query Match 84.2%; Score 32; DB 1; Length 40;  
Best Local Similarity 80.0%; Pred. No. 19;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 WLWVI 5  
DB 11 MWVMI 15  
  
-----  
RESULT 11  
MERT\_SHEPU STANDARD: PRT; 115 AA.  
ID MERT\_SHEPU  
AC Q54462;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Mercuric transport protein (Mercury ion transport protein).  
GN MERT.  
OS Shewanella putrefaciens (Pseudomonas putrefaciens).  
OC Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;  
CC Shewanella.  
CC NCBI\_TaxID=24;  
CC (1)  
CC SEQUENCE FROM N.A.  
CC MEDLINE-97309939; PubMed-9167257;  
CC RA Osborn A.M., Bruce R.D., Strike P., Ritchie D.A.;  
CC "Distribution, diversity and evolution of the bacterial mercury  
CC resistance (mer) operon."  
CC FEMS Microbiol. Rev. 19:239-262(1997).  
CC -1- FUNCTION: INVOLVED IN MERCURIC TRANSPORT. PASSES A HG(2+) ION  
CC FROM THE PERIPLASMIC MERT PROTEIN TO THE MERCURIC REDUCTASE  
CC (MERT).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
CC (BY SIMILARITY).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z49196; CAAB9054.1; -;  
CC DR InterPro: IPR003457; Transprt\_Mert.  
CC DR Pfam: PF02411; MERT; 1.  
CC KM Transport; Mercuric resistance; Inner membrane; Mercury; Plasmid;  
CC Transmembrane.  
CC KW TRANSMEM 10 30 POTENTIAL.  
CC FT TRANSMEM 45 65 POTENTIAL.  
CC FT TRANSMEM 92 112 POTENTIAL.  
CC FT METAL 23 23 HG(2+) (BY SIMILARITY).  
CC FT METAL 24 24 HG(2+) (BY SIMILARITY).  
CC FT METAL 75 75 HG(2+) (BY SIMILARITY).  
CC FT METAL 81 81 HG(2+) (BY SIMILARITY).  
CC SQ SEQUENCE 115 AA: 12555 MW: C4ADC83CD389074E CRC64;  
  
Query Match 84.2%; Score 32; DB 1; Length 115;  
Best Local Similarity 66.7%; Pred. No. 48;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLWVI 6  
DB 110 WLWVFA 115  
  
-----  
RESULT 12  
ERD2\_PETHY STANDARD: PRT; 215 AA.  
ID ERD2\_PETHY  
AC Q9ZTN2;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE ER lumen protein retaining receptor (HDEL receptor) (PGP169-12).  
GN ERD2.  
OS Petunia hybrida (Petunia).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Petunia.  
CC NCBI\_TaxID=4102;  
CC (1)  
CC SEQUENCE FROM N.A.  
CC MEDLINE-20317212; PubMed-10859200;  
CC Guyon V.N., Astwood J.D., Garner E.C., Dunker A.K., Taylor L.P.;  
CC "Isolation and characterization of cDNAs expressed in the early  
CC stages of flavonol-induced pollen germination in petunia."  
CC Plant Physiol. 123:699-710(2000).  
CC -1- FUNCTION: REQUIRED FOR THE RETENTION OF LUMINAL ENDOPLASMIC  
CC RETICULUM PROTEINS, DETERMINES THE SPECIFICITY OF THE LUMINAL ER  
CC PROTEIN RETENTION SYSTEM. ALSO REQUIRED FOR NORMAL VESICULAR  
CC TRAFFIC THROUGH THE GOLGI. THIS RECEPTOR RECOGNIZES H-D-E-L (BY  
CC SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC (BY SIMILARITY).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF049922; AAD02548.1; -;  
CC DR InterPro: IPR000133; Eretl\_receptor.  
CC DR Pfam: PF00810; ER\_lumen\_recept; 1.  
CC DR PRINTS: PR00660; ERLUMENR.  
CC DR PRODOM: PD005774; Eretl\_receptor; 1.  
CC DR PROSITE: PS00951; ER\_LUMEN\_RECEPTOR\_1;  
CC DR PROSITE: PS00952; ER\_LUMEN\_RECEPTOR\_2; 1.  
CC KW Endoplasmic reticulum; Transmembrane; Protein transport; Receptor.  
CC FT DOMAIN 1 2 LUMENAL (POTENTIAL).  
CC FT TRANSMEM 3 21 POTENTIAL.  
CC FT DOMAIN 22 35 CYTOPLASMIC (POTENTIAL).  
CC FT TRANSMEM 36 53 POTENTIAL.  
CC FT DOMAIN 54 61 LUMENAL (POTENTIAL).  
CC FT TRANSMEM 62 82 POTENTIAL.  
CC FT DOMAIN 83 98 CYTOPLASMIC (POTENTIAL).  
CC FT TRANSMEM 99 112 POTENTIAL.  
CC FT DOMAIN 113 119 LUMENAL (POTENTIAL).  
CC FT TRANSMEM 120 139 POTENTIAL.  
CC FT DOMAIN 140 151 CYTOPLASMIC (POTENTIAL).  
CC FT TRANSMEM 152 170 POTENTIAL.  
CC FT DOMAIN 171 181 LUMENAL (POTENTIAL).  
CC FT TRANSMEM 182 202 POTENTIAL.  
CC FT DOMAIN 203 215 CYTOPLASMIC (POTENTIAL).  
CC SQ SEQUENCE 215 AA: 25602 MW: 7A6E4EF4F01CF65 CRC64;  
  
Query Match 84.2%; Score 32; DB 1; Length 215;  
Best Local Similarity 66.7%; Pred. No. 83;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 181 WITWIA 186

## RESULT 13

AOPZ\_BRUBA STANDARD: PRT; 228 AA.

AC 09LA79;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Aquaporin 2 (Aquaporin X).

GN AOPZ OR AOPX.

OS Brucella abortus.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Brucellaceae; Brucella.

OX NCBI\_TaxID=235;

RN [1]

SEQUENCE FROM N.A.

STRAIN=2308;

MEDLINE=20553188; PubMed=11101683;

RA Rodriguez M.C., Proger A., Rolland J.-P., Thomas D., Aguero J.,

RA Delamarque C., Garcia-Lobo J.M.,

RT "A functional water channel protein in the pathogenic bacterium

Brucella abortus."

RL Microbiology 146:3251-3257(2000).

CC -1- FUNCTION: Transport of water across the membrane. Possibly

involved in the adaptation to variation in intracellular pH or

osmolality.

CC -1- SUBUNIT: Homotrimer (By similarity).

CC -1- SIMILARITY: BELONGS TO THE MIP/AQUAPORIN FAMILY (TC 1.A.8).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AF148066; AAF3105.1; -

DR HSSP: P29972; 1H61.

DR InterPro: IPR000425; MIP\_family.

DR Pfam: PF00230; MIP: 1.

DR PRINTS: PR00783; MINTRENSICP.

DR ProDom: PD000295; MIP\_family: 1.

DR TIGRFAMs: TIGR00861; MIP: 1.

DR PROSITE: PS00221; MIP: 1.

KW Transport; Transmembrane; Inner membrane; Repeat.

FT TRANSMEM 7 29 POTENTIAL.

FT TRANSMEM 34 56 POTENTIAL.

FT TRANSMEM 81 103 POTENTIAL.

FT TRANSMEM 127 149 POTENTIAL.

FT TRANSMEM 154 176 POTENTIAL.

FT TRANSMEM 203 225 POTENTIAL.

FT REPEAT 63 65 NPA 1.

FT REPEAT 184 186 NPA 2.

FT SITE 20 20 INVOLVED IN TETRAMERIZATION OR STABILITY

OF THE TETRAMER (BY SIMILARITY).

FT DOMAIN 92 97 POLY-ALA.

FT SEQUENCE 228 AA; 23145 MW; 23664C39E4DD44A CRC64;

Query Match Score 32; DB 1; Length 228;

Best Local Similarity 66.7%; Pred. No. 88;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLVWIA 6

DB 204 WLVWIA 209

RESULT 14

AOPZ\_BRUME STANDARD: PRT; 228 AA.

AC 09L772;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Aquaporin 2.

GN AOPZ OR BME10070.

OS Brucella melitensis.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Brucellaceae; Brucella.

OX NCBI\_TaxID=29459;

RN [1]

SEQUENCE FROM N.A.

STRAIN=133;

RA Hernandez-Castro R., Verdugo-Rodriguez A., Gutierrez-Pabello J.A.,

RA Suarez-Guemes F.;

RT "Cloning and nucleotide sequence of aquaporin gene of Brucella

melitensis."

RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

SEQUENCE FROM N.A.

RC STRAIN=16M / ATCC 23456 / Biotype 1;

RC MEDLINE=20020109; PubMed=11756688;

RA DelVecchio V.G., Kaputal V., Redkar R.J., Patra G., Muijer C., Los T.,

RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Reznik G.,

RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,

RA Selkov E., Elizer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,

RA Haselkorn R., Kyriades N., Overbeek R.;

RT "The genome sequence of the facultative intracellular pathogen

Brucella melitensis."

RT Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).

CC -1- FUNCTION: Transport of water across the membrane. Possibly

involved in the adaptation to variation in intracellular pH or

osmolality (By similarity).

CC -1- SUBUNIT: Homotrimer (By similarity).

CC -1- SIMILARITY: BELONGS TO THE MIP/AQUAPORIN FAMILY (TC 1.A.8).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AF26624; AAF36396.1; -

DR HSSP: P29972; 1H61.

DR InterPro: IPR000425; MIP\_family.

DR Pfam: PF00230; MIP: 1.

DR PRINTS: PR00783; MINTRENSICP.

DR ProDom: PD000295; MIP\_family: 1.

DR TIGRFAMs: TIGR00861; MIP: 1.

DR PROSITE: PS00221; MIP: 1.

KW Transport; Transmembrane; Inner membrane; Repeat; Complete proteome.

FT TRANSMEM 7 29 POTENTIAL.

FT TRANSMEM 34 56 POTENTIAL.

FT TRANSMEM 81 103 POTENTIAL.

FT TRANSMEM 127 149 POTENTIAL.

FT TRANSMEM 154 176 POTENTIAL.

FT TRANSMEM 203 225 POTENTIAL.

FT REPEAT 63 65 NPA 1.

FT REPEAT 184 186 NPA 2.

FT SITE 20 20 INVOLVED IN TETRAMERIZATION OR STABILITY

OF THE TETRAMER (BY SIMILARITY).

FT DOMAIN 92 97 POLY-ALA.

FT SEQUENCE 228 AA; 23189 MW; 64282C39E4DD4109 CRC64;

Query Match Score 32; DB 1; Length 228;

Best Local Similarity 66.7%; Pred. No. 88;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLWMA 6  
1111  
DB 204 WLFWA 209

RESULT 15

Y836\_AQUAE  
ID Y836\_AQUAE STANDARD; PRT; 232 AA.  
AC 067008;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein AQ\_836.

OS Aquifex aeolicus.

OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;

OC Aquifex.

NCBI\_TaxID=63363;

[1]  
SEQUENCE FROM N.A.

STRAIN=VF5;

RX MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,

RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,

RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,

RT "The complete genome of the hyperthermophilic bacterium Aquifex

aeolicus."

RL Nature 392:353-358(1998).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: AE00709; AAC06970.1; -

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 4 24 POTENTIAL.

FT TRANSMEM 42 62 POTENTIAL.

FT TRANSMEM 100 120 POTENTIAL.

FT TRANSMEM 145 165 POTENTIAL.

FT TRANSMEM 171 191 POTENTIAL.

SQ SEQUENCE 232 AA; 26317 MW; B934B56FEA35B183 CRC64;

Query Match

Best Local Similarity 84.2%; Score 32; DB 1; Length 232;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWMA 5

1111

DB 183 WLWML 187

Search completed: January 3, 2003, 15:29:24

Job time : 6.18182 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:24:35 ; Search time 20.3182 seconds  
(without alignments)  
60.846 Million cell updates/sec

Title: US-09-543-188a-27

Perfect score: 38

Sequence: 1 WLWVIA 6

Scoring table: BLOSUM62

Searched: 671580 seqs, 206047115 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21.\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*  
15: sp-virus:\*  
16: sp-bacteriap:\*  
17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	100.0	371	16	Q92D92
2	38	100.0	371	16	Q8Y8H8
3	38	100.0	598	16	Q8Y8H8
4	37	97.4	719	16	Q922H9
5	36	94.7	438	16	Q8XPT4
6	36	94.7	651	16	Q8YLI2
7	35	92.1	234	16	Q8CK00
8	35	92.1	358	5	Q967S2
9	35	92.1	418	16	Q8Y0V2
10	35	92.1	463	5	Q967S1
11	35	92.1	468	16	Q8Z9F2
12	35	92.1	468	16	Q9L4I1
13	35	92.1	536	5	Q19337
14	35	92.1	555	5	Q906X2
15	35	92.1	556	5	Q967S3
16	35	92.1	585	5	O45298

17	34	89.5	61	2	Q9AML5
18	34	89.5	162	8	Q942A9
19	34	89.5	162	8	Q94292
20	34	89.5	168	9	Q8SD14
21	34	89.5	247	11	Q8RI89
22	34	89.5	285	16	Q67817
23	34	89.5	296	3	Q32486
24	34	89.5	316	2	Q07450
25	34	89.5	326	17	Q8U1U9
26	34	89.5	384	16	Q69496
27	34	89.5	453	3	Q02922
28	34	89.5	587	10	Q92T77
29	34	89.5	602	5	Q8SMH6
30	34	89.5	703	16	Q927C0
31	34	89.5	714	16	Q9K1S6
32	34	89.5	1006	3	Q07324
33	34	89.5	1040	3	Q94147
34	34	89.5	1394	3	Q08409
35	33	86.8	162	8	Q942B0
36	33	86.8	162	8	Q942A7
37	33	86.8	162	8	Q942A6
38	33	86.8	162	8	Q942A5
39	33	86.8	162	8	Q942A2
40	33	86.8	162	8	Q94298
41	33	86.8	162	8	Q94298
42	33	86.8	162	8	Q94296
43	33	86.8	162	8	Q94YV3
44	33	86.8	170	17	Q8U429
45	33	86.8	171	10	Q9LV89

## ALIGNMENTS

RESULT 1  
ID Q92D92 PRELIMINARY: PRT: 371 AA.  
AC Q92D92:  
DT 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
DE Putative membrane protein.  
GN LINO925.  
OS Listeria innocua.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Listeriaceae; Listeria.  
OX NCBI\_TaxID=1642;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CLIP 11262 / SEROVAR 6A;  
RX PubMed=11679669;  
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deboux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,  
RA Ertan K.-D., Esbl H., Garcia-del Portillo F., Garrido P.,  
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkai G.,  
RA Madueno E., Maltournam A., Mata Vicente J., Ng E., Nedjati R.,  
RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schleuter T., Simoes N., Tierrez A.,  
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,  
RT "Comparative genomics of Listeria species.";  
RL Science 294:849-852(2001).  
DR EMBL; AL596166; CAC96157.1; -.  
DR L1stlist; LINO0925; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 371 AA: 41360 MW: 46FE4A30E917C242 CRC64;  
Query Match 100.0%; Score 38; DB 16; Length 371;  
Best Local Similarity 100.0%; Pred. No. 89;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWVIA 6  
 |||||  
 DB 313 WLWVIA 318

## RESULT 2

O8Y8H8

AC 08Y8H8 PRELIMINARY: PRT: 371 AA.

DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

GN LMO0925.

OS Listeria monocytogenes.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Listeriaceae; Listeria.

OX NCBI\_TaxID=1639;

RN [1]

SEQUENCE FROM N.A.

STRAIN-EGD-E / SEROVAR 1/2A:

MEDLINE-21537279; PubMed11679669;

Glasier P., Frangeul L., Buchrieser C., Rusnok C., Amend A.,

Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,

Charbit A., Chetouani F., Couve E., de Daruvar A., Deboux P.,

Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,

Ertlan K.-D., Fsihl H., Garcia-del Portillo F., Garrido P.,

Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkai G.,

Madueno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,

Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,

Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,

Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;

RT "Comparative genomics of Listeria species.";

RL Science 294:849-852(2001).

DR EMBL: AL591977; CAC99003.1; -.

DR Lifestlist; LMO00925; -.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 371 AA; 41252 MW; 54E7E86545CA62E CRC64;

Query Match 100.0%; Score 38; DB 16; Length 371;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWVIA 6

|||||

DB 313 WLWVIA 318

## RESULT 3

O8Y063

AC 08Y063 PRELIMINARY: PRT: 598 AA.

DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Probable inner membrane transmembrane protein.

GN RSC1181 OR RS05713.

OS Ralstonia solanacearum (Pseudomonas solanacearum).

OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;

OC Ralstonia.

OX NCBI\_TaxID=305;

RN [1]

SEQUENCE FROM N.A.

STRAIN-EM11000;

MEDLINE-21681879; PubMed11823852;

Arlat M., Billault A., Brottier P., Camus J.C., Cattoilco L.,

Chandler M., Choise N., Claudel-Bernard C., Cunac S., Demange N.,

Gaspin C., Layle M., Moisan A., Robert C., Saurin W., Schlex T.,

Sigler P., Thebaud P., Whalen M., Wincker P., Levy M.,

Weissenbach J., Boucher C.A.;

RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";

RL Nature 415:497-502(2002).  
 DR EMBL: AL646063; CAD14883.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 598 AA; 66422 MW; 8798BC65EA5C6646 CRC64;

Query Match 100.0%; Score 38; DB 16; Length 598;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWVIA 6

|||||

DB 393 WLWVIA 398

## RESULT 4

O9Z2H9

AC 09Z2H9 PRELIMINARY: PRT: 719 AA.

DT 01-DEC-2001 (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Hypothetical protein RA0506.

GN RA0506 OR SMA0937.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Plasmid psyma (megaplasmid 1).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI\_TaxID=382;

RN [1]

SEQUENCE FROM N.A.

STRAIN-1021;

MEDLINE-21396509; PubMed11481432;

Barrett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,

Barloy-Hubler F., Bowser L., Capela D., Galburt F., Gouzy J.,

Gurjal M., Hong A., Huzar L., Hyman R.W., Kahn D., Kahn M.L.,

Kalan S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,

Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;

RT "Nucleotide sequence and predicted functions of the entire

Sinorhizobium meliloti psyma megaplasmid.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).

DR EMBL: AE007241; AAK65164.1; -.

DR InterPro: IPR001880; MSion\_channel.

DR InterPro: IPR000504; RNA\_rec\_mot.

DR Pfam: PF00924; MS\_channel.1.

DR PROSITE: PS00030; RRM\_RNP.1; UNKNOWN.1.

KW Plasmid; Hypothetical protein; Complete proteome.

SQ SEQUENCE 719 AA; 77245 MW; 045DBCB16917B9 CRC64;

Query Match 97.4%; Score 37; DB 16; Length 719;  
 Best Local Similarity 83.3%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWVIA 6

|||||

DB 333 WLWVIA 338

## RESULT 5

O8XPT4

AC 08XPT4 PRELIMINARY: PRT: 438 AA.

DT 01-MAR-2002 (TREMBlrel. 20, Created)

DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE regulator protein (EC 2.7.-.-).

GN RSP1553 OR RS02109.

OS Ralstonia solanacearum (Pseudomonas solanacearum).

OC Plasmid megaplasmid.

OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;

OC Ralstonia.

OX NCBI\_TaxID=305;

RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-GM11000;  
RX MEDLINE-21681879; PubMed-11823852;  
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Manjencot S.,  
RA Ariat M., Billault A., Brottier P., Camus J.C., Catolico L.,  
RA Chandler M., Chouane N., Claudel-Renard C., Cunnac S., Demange N.,  
RA Gaspin C., Layve M., Moisan A., Robert C., Saurin W., Schlex T.,  
RA Siglier P., Tebault P., Whalen M., Wincker P., Levy M.,  
RA Weissenbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*.";  
RL Nature 415:497-502(2002).  
DR EMBL: AL646085; CAD18704.1; -;  
DR InterPro: IPR003594; ATPbind\_ATPase.  
DR InterPro: IPR004358; Bact\_sens\_pr\_C.  
DR InterPro: IPR003660; HAMF.  
DR InterPro: IPR004359; HIS\_KIN\_s19.  
DR Pfam: PF00672; HAMF; 1.  
DR Pfam: PF02518; HATPase\_C; 1.  
DR Pfam: PF00512; signal; 1.  
DR PRINTS: PR00344; BCTRLSENSOR.  
DR SMART: SM00387; HATPase\_C; 1.  
DR SMART: SM00388; HSKA; 1.  
DR PROSITE: PS50109; HIS\_KIN; 1.  
DR Transferrase; Plasmid; Complete proteome.  
SQ SEQUENCE 438 AA; 46589 MW; A28879BB7E0B074D CRC64;

Query Match 94.7%; Score 36; DB 16; Length 438;  
Best Local Similarity 83.3%; Pred. No. 2.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWVIA 6  
Db 155 WLWVIA 160

RESULT 6  
O8YL12 PRELIMINARY; PRT; 651 AA.  
AC O8YL12;  
DT 01-MAR-2002 (TREMblrel. 20, Created)  
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)  
DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
DE Hypothetical protein Alr7126.  
GN ALR7126.  
OS Anabaena sp. (strain PCC 7120).  
OC Plasmid PCC7120alpha.  
NCBI\_Taxid=103690;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-21595285; PubMed-11759840;  
RA Kaneo T., Nakamura Y., Wolk G.P., Kuritz T., Sasamoto S.,  
RA Watanabe A., Iritiguchi M., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kohara M., Matsuno M., Matsuno A., Muraki A.,  
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
RA Yasuda M., Tabata S.;  
RT "Complete genomic sequence of the filamentous nitrogen-fixing  
cyanobacterium *Anabaena* sp. strain PCC 7120.";  
RL DNA Res. 8:205-213(2001).  
DR EMBL: AF003600; BAB78210.1; -;  
KW Plasmid; Hypothetical protein; Complete proteome.  
SQ SEQUENCE 651 AA; 74404 MW; 6AB753997CD2C809 CRC64;

Query Match 94.7%; Score 36; DB 16; Length 651;  
Best Local Similarity 83.3%; Pred. No. 3.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWVIA 6  
Db 369 WLWVIA 374

RESULT 7  
O9CK00 PRELIMINARY; PRT; 234 AA.  
AC O9CK00;  
DT 01-JUN-2001 (TREMblrel. 17, Created)  
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE Hypothetical protein PM188.  
GN PM188.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Pasteurella.  
NCBI\_Taxid=747;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Pm70.  
RX MEDLINE-21145866; PubMed-11248100;  
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;  
RT "Complete genomic sequence of *Pasteurella multocida* Pm70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
DR EMBL: AF006221; AAK03922.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 234 AA; 26264 MW; 0B31A94DD2AA2F9 CRC64;

Query Match 92.1%; Score 35; DB 16; Length 234;  
Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWVIA 6  
Db 180 WLWVIA 185

RESULT 8  
O967S2 PRELIMINARY; PRT; 358 AA.  
AC O967S2;  
DT 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
DE Nuclear receptor nhr-7B.  
GN NHR-7.  
OS Dirofilaria immitis (Canine heartworm).  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Splurrida; Filarioidea;  
OC Onchocercidae; Dirofilaria.  
NCBI\_Taxid=6287;  
[1]  
RP SEQUENCE FROM N.A.  
RA Crossgrove K., Laudet V., Maina C.V.;  
RT "Dirofilaria immitis encodes dlnhr-7, a putative homolog of the  
Drosophila ecdysone-regulated E78 gene.";  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
DR EMBL: AF367207; AAK54128.1; -;  
DR InterPro: IPR000536; Hormone\_rec\_1lg.  
DR InterPro: IPR001628; znf\_C4steroid.  
DR Pfam: PF00104; hormone\_rec; 1.  
DR Pfam: PF00105; zf-C4; 1.  
DR Prodom: PD000035; znf\_C4steroid; 1.  
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;  
KW Zinc-finger.  
SQ SEQUENCE 358 AA; 41669 MW; 2C01ED1D96C86C9E CRC64;

Query Match 92.1%; Score 35; DB 5; Length 358;  
Best Local Similarity 83.3%; Pred. No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWVIA 6  
Db 166 WLWVIA 171

```
RESULT 9
Q8Y0V2 PRELIMINARY; PRT; 418 AA.
AC Q8Y0V2;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Probable integral membrane protease transmembrane protein
DE (EC 3.4.-).
GN RSC0941 OR RS04465.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM1100;
MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
  Arlat M., Billault A., Brottier P., Camus J.C., Catolico L.,
  Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
  Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
  Weissenbach J., Boucher C.A.;
RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502(2002).
DR EMBL: AL646061; CAD14643.1; -
DR InterPro: IPR001915; Peptidase_M48.
DR Pfam: PF01435; Peptidase_M48; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 418 AA; 46610 MW; 8ACCD9903C85D52 CRC64;

Query Match
Best Local Similarity 92.1%; Score 35; DB 16; Length 418;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6
Db 180 WLWVMA 185

RESULT 10
Q967S1 PRELIMINARY; PRT; 463 AA.
AC Q967S1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Nuclear receptor nhr-7C.
DE NHR-7.
OS Drosophila immitis (Canine heartworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Drosophila.
OX NCBI_TaxID=6287;
RN [1]
RP SEQUENCE FROM N.A.
RA "Drosophila immitis encodes dlnhr-7, a putative homolog of the
  Drosophila ecdysone-regulated E78 gene.";
  Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
  CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
  DE -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
  DE EMBL: AF367208; AK54129.1; -
  DR InterPro: IPR000536; Hormone_rec_119.
  DR InterPro: IPR001628; ZnF_C4steroid.
  DR Pfam: PF00104; hormone_rec; 1.
  DR ProDom: PD000035; znf_C4steroid; 1.
  DR PROSITE: PS00031; NUCLEAR_RECEPTOR; UNKNOWN_1.
  KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
  KW Zinc-finger.
  SQ SEQUENCE 463 AA; 52253 MW; A3430D0FAC35DF8E CRC64;
```

```
Query Match
Best Local Similarity 92.1%; Score 35; DB 5; Length 463;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6
Db 360 WLWVIS 365

RESULT 11
Q829F2 PRELIMINARY; PRT; 468 AA.
AC Q829F2;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative symporter.
DE SY0171.
GN Salmonella typhi.
OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
  RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
  RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
  RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
  RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
  RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
  RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
  RA Whitehead S., Barrall B.G.;
RA "Complete genome sequence of a multiple drug resistant Salmonella
  RT enterica serovar Typhi CT18.";
  RL Nature 413:848-852(2001).
  DR EMBL: AL627265; CAD01308.1; -
  DR InterPro: IPR001092; HLH_basic.
  DR InterPro: IPR001927; Na/Gal_symp.
  DR TIGRPFAM: TIGR00792; gph; 1.
  DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
  KW Hypothetical protein; Complete proteome.
  SQ SEQUENCE 468 AA; 52589 MW; 883F0DF7E8D3A87 CRC64;

Query Match
Best Local Similarity 92.1%; Score 35; DB 16; Length 468;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6
Db 82 WLWVIA 87

RESULT 12
Q9L4I1 PRELIMINARY; PRT; 468 AA.
AC Q9L4I1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 52.4 kDa protein (Putative permease of the
  DE Na+/galactoside symporter family).
  DE SYW0149.
  GN Salmonella typhimurium.
  OS Salmonella typhimurium.
  OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
  OC Salmonella.
  OX NCBI_TaxID=602;
  RN [1]
  RP SEQUENCE FROM N.A.
  RC STRAIN=SL1344;
  RA Cano D., Casadesus J., Garcia-del Portillo F.;
  RA "Characterization of a Salmonella-specific region located between ampe
```



```

RT and arcp genes."
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Gireal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2."
RL Nature 413:852-856(2001).
DR EMBL: AJ242516; CAB89838.1;
DR EMBL: AE008700; AAL19113.1;
DR InterPro: IPR001927; HLH_basic.
DR TIGRfams: TIGR00792; gph; 1.
PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
SQ Hypothetical protein; Complete proteome.
SQ SEQUENCE 468 AA; 52443 MW; 5545B43CB3EAB0B3A CRC64;

Query Match
Best Local Similarity 92.1%; Score 35; DB 16; Length 468;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWIA 6
DB 82 WLWIA 87

RESULT 13
Q19337 PRELIMINARY; PRT; 536 AA.
AC Q19337;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F11A10.5 protein.
GN F11A10.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC Smye R.;
RT Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: Z68297; CAA92595.1;
SQ SEQUENCE 536 AA; 60611 MW; 0449018148FBC181 CRC64;

Query Match
Best Local Similarity 92.1%; Score 35; DB 5; Length 536;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWIA 6
DB 9 WLWIA 14

RESULT 14
Q9U6X2 PRELIMINARY; PRT; 555 AA.
AC Q9U6X2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Nuclear hormone receptor.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20058753; PubMed=10593180;
RA Unnasch T.R., Bradley J., Beauchamp J., Tuan R., Kennedy M.W.;
RT "Characterization of a putative nuclear receptor from Onchocerca
RT volvulus."
RL Mol. Biochem. Parasitol. 104:259-269(1999).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL: AF170551; AAF06997.1;
DR HSSP: P20393; 1A6Y.
DR InterPro: IPR000536; Hormone_rec.119.
DR InterPro: IPR001723; Steroid_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; hormone_rec; 1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR00398; STRDHORMONER.
DR PRINTS: PR00047; STROIDFINGER.
DR ProDom: PD000035; Znf_C4steroid; 1.
DR SMART: SM00430; HOL1.1.
DR SMART: SM00399; Znf_C4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 555 AA; 62847 MW; E813CB9F20C94087 CRC64;

Query Match
Best Local Similarity 92.1%; Score 35; DB 5; Length 555;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWIA 6
DB 363 WLWIA 368

RESULT 15
Q967S3 PRELIMINARY; PRT; 556 AA.
AC Q967S3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Nuclear receptor nhr-7A.
GN NHR-7.
OS Dirofilaria immitis (Canine heartworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Dirofilaria.
OX NCBI_TaxID=6287;
RN [1]
RP SEQUENCE FROM N.A.
RC Crossgrove K., Laudet V., Maina C.V.;
RA "Dirofilaria immitis encodes dinhr-7, a putative homolog of the
RT drosophila ecdysone-regulated E78 gene."
RT Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL: AF367206; AAK54127.1;
DR InterPro: IPR000536; Hormone_rec.119.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; hormone_rec; 1.
DR Pfam: PF00105; zf-C4; 1.
DR ProDom: PD000035; Znf_C4steroid; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 556 AA; 63342 MW; 2F7995FAFB5126F CRC64;

```

Query Match 92.1%; Score 35; DB 5; Length 556;  
Best Local Similarity 83.3%; Pred. No. 3.8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0;  
QY 1 WLWIA 6  
|||||:  
Db 364 WLWIS 369

Search completed: January 3, 2003, 15:32:04  
Job time : 21.3182 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:16 ; Search time 9 seconds  
(without alignments)  
19.615 Million cell updates/sec

Title: US-09-543-188a-27

Perfect score: 38

Sequence: 1 MLWIA 6

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Database: 262574 seqs, 29422922 residues

Number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents, AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/Dackfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	86.8	362	1	US-08-118-270-16
2	33	86.8	362	5	PCT-US93-08528-16
3	33	86.8	572	6	5200183-5
4	32	84.2	75	4	US-08-469-260A-478
5	32	84.2	139	1	US-08-137-117D-35
6	32	84.2	139	2	US-08-436-717-35
7	32	84.2	266	6	5223610-5
8	32	84.2	269	6	5244657-5
9	32	84.2	269	6	5433945-5
10	32	84.2	837	4	US-09-122-126B-2
11	31	81.6	50	4	US-08-469-260A-534
12	31	81.6	377	4	US-09-316-080-5
13	31	81.6	380	4	US-09-316-080-2
14	31	81.6	438	2	US-08-952-365-2
15	31	81.6	705	2	US-08-663-566A-19
16	31	81.6	705	2	US-08-023-610-19
17	31	81.6	705	2	US-08-288-065A-19
18	31	81.6	705	2	US-08-362-240A-19
19	31	81.6	705	5	PCT-US95-10245-19
20	31	81.6	1162	2	US-08-663-566A-15
21	31	81.6	1162	2	US-08-023-610-15
22	31	81.6	1162	2	US-08-288-065A-15
23	31	81.6	1162	2	US-08-362-240A-15
24	31	81.6	1162	5	PCT-US95-10245-15
25	30	78.9	24	2	US-08-470-419-2
26	30	78.9	24	2	US-08-761-828-2
27	30	78.9	24	2	US-08-290-105-2

28	30	78.9	24	3	US-08-776-949-2	Sequence 2, Appl
29	30	78.9	24	3	US-08-482-810-2	Sequence 2, Appl
30	30	78.9	24	4	US-09-027-955-2	Sequence 2, Appl
31	30	78.9	24	4	US-09-636-805-2	Sequence 2, Appl
32	30	78.9	24	4	US-09-258-128-2	Sequence 2, Appl
33	30	78.9	52	2	US-08-470-419-4	Sequence 4, Appl
34	30	78.9	52	2	US-08-761-828-4	Sequence 4, Appl
35	30	78.9	52	2	US-08-290-105-4	Sequence 4, Appl
36	30	78.9	52	3	US-08-776-949-4	Sequence 4, Appl
37	30	78.9	52	3	US-08-482-810-4	Sequence 4, Appl
38	30	78.9	52	4	US-09-027-955-4	Sequence 4, Appl
39	30	78.9	52	4	US-09-636-805-4	Sequence 4, Appl
40	30	78.9	143	2	US-09-258-128-4	Sequence 4, Appl
41	30	78.9	143	3	US-08-834-033A-112	Sequence 11, Appl
42	30	78.9	143	3	US-08-834-033A-112	Sequence 11, Appl
43	30	78.9	143	4	US-09-363-574-11	Sequence 11, Appl
44	30	78.9	143	4	US-09-363-574-11	Sequence 11, Appl
45	30	78.9	145	5	PCT-US94-01149-69	Sequence 69, Appl

#### ALIGNMENTS

RESULT 1  
US-08-118-270-16  
; Sequence 16, Application US/08118270  
; Patent No. 5508384  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Randall B.  
; APPLICANT: Schuster, David I.  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
; NUMBER OF SEQUENCES: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEWMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/118,270  
; FILING DATE: 09-SEP-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/943,236  
; FILING DATE: 10-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Townsend, Kevin G.  
; REGISTRATION NUMBER: 34,033  
; REFERENCE/DOCKET NUMBER: MURPHY-2A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; TELEX: 248633  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 362 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-118-270-16  
Query Match 86.8%; Score 33; DB 1; Length 362;  
Best Local Similarity 80.0%; Pred. No. 3.7e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLWIA 5

Db 120 WLVWV 124

||||:

## RESULT 2

PCT-US93-08528-16  
; Sequence 16, Application PC/TUS9308528  
; GENERAL INFORMATION:  
; APPLICANT: New York University  
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN  
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF  
; NUMBER OF SEQUENCES: 348  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/08528  
; FILING DATE: 09-SEP-1993  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 07/943,236  
; FILING DATE: 10-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Townsend, Kevin G.  
; REGISTRATION NUMBER: 34,033  
; REFERENCE/DOCKET NUMBER: MURPHY-2 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; TELEX: 248633  
; INFORMATION FOR SEQ. ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 362 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; PCT-US93-08528-16

## Query Match

Best Local Similarity 86.8%; Score 33; DB 5; Length 362;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLVWV 5

Db 120 WLVWV 124

## RESULT 3

5200183-5  
; Patent No. 5200183  
; APPLICANT: TANG, JORDAN J.N.; WANG, CHI-SUN  
; TITLE OF INVENTION: RECOMBINANT BILE SALT ACTIVATED LIPASES  
; NUMBER OF SEQUENCES: 22  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/537,426  
; FILING DATE: 12-JUN-1990  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 504,635  
; FILING DATE: 04-APR-1990  
; APPLICATION NUMBER: 122,410  
; FILING DATE: 19-NOV-1987  
; SEQ. ID NO: 5:  
; LENGTH: 572  
5200183-5

## Query Match

Best Local Similarity 86.8%; Score 33; DB 6; Length 572;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLVWV 5

Db 107 WLVWV 111

## RESULT 4

US-08-469-260A-478  
; Sequence 478, Application US/08469260A  
; Patent No. 6451578  
; GENERAL INFORMATION:  
; APPLICANT: JOHN N. SIMONS  
; APPLICANT: TAMI J. PILOT-MATIAS  
; APPLICANT: GEORGE J. DAWSON  
; APPLICANT: GEORGE G. SCHLAUDER  
; APPLICANT: SURESH M. DESAI  
; APPLICANT: THOMAS P. LEARY  
; APPLICANT: ANTHONY SCOTT MUEHROFF  
; APPLICANT: JAMES C. ERKER  
; APPLICANT: SHERI L. BUIEK  
; APPLICANT: ISA K. MUSHAMMAR  
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS  
; NUMBER OF SEQUENCES: 716  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D  
; STREET: 100 ABBOTT PARK ROAD  
; CITY: ABBOTT PARK  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,260A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US/08/424,550  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: POREMSKI, PRISCILLA E.  
; REGISTRATION NUMBER: 33,207  
; REFERENCE/DOCKET NUMBER: 5527.PC.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 708-938-2623  
; TELEFAX: 708-937-6365  
; INFORMATION FOR SEQ. ID NO: 478:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 75 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-469-260A-478

## Query Match

Best Local Similarity 84.2%; Score 32; DB 4; Length 75;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLVWV 6

Db 2 WLVWV 7

## RESULT 5

US-08-137-117D-35

Sequence 35, Application US/08137117D  
Patent No. 5795965  
GENERAL INFORMATION:  
APPLICANT: TSUCHIYA, Masayuki  
APPLICANT: SATO, Koh  
APPLICANT: BENDIG, Mary  
APPLICANT: JONES, Steven  
APPLICANT: SALDANHA, Jose  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
INTERLEUKIN-6 RECEPTOR  
NUMBER OF SEQUENCES: 158  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,117D  
FILING DATE: 20-DEC-1993  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/JP92/00544  
FILING DATE: 24-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-32084  
FILING DATE: 19-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 3-95476  
FILING DATE: 25-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ. ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-137-117D-35

Query Match 84.2%; Score 32; DB 1; Length 139;  
Best Local Similarity 80.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVI 5  
DB 52 WLWVI 56

RESULT 6  
US-08-436-717-35  
Sequence 35, Application US/08436717  
Patent No. 581790  
GENERAL INFORMATION:  
APPLICANT: TSUCHIYA, Masayuki  
APPLICANT: SATO, Koh  
APPLICANT: BENDIG, Mary  
APPLICANT: JONES, Steven  
APPLICANT: SALDANHA, Jose  
TITLE OF INVENTION: RESHAPED HUMAN ANTIBODY TO HUMAN  
INTERLEUKIN-6 RECEPTOR

NUMBER OF SEQUENCES: 158  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,717  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,117  
FILING DATE: 20-DEC-1993  
APPLICATION NUMBER: WO PCT/JP92/00544  
FILING DATE: 24-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 4-32084  
FILING DATE: 19-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 3-95476  
FILING DATE: 25-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: WEGNER, Harold C.  
REGISTRATION NUMBER: 25,258  
REFERENCE/DOCKET NUMBER: 53466/126/AAOK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ. ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-436-717-35

Query Match 84.2%; Score 32; DB 2; Length 139;  
Best Local Similarity 80.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVI 5  
DB 52 WLWVI 56

RESULT 7  
5223610-5  
Patent No. 5223610  
APPLICANT: Burton, Frank H.; Sutcliffe, Gregor  
TITLE OF INVENTION: CHOLERA TOXIN GENE REGULATED BY GROWTH  
HORMONE PROMOTER  
NUMBER OF SEQUENCES: 18  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/528,852  
FILING DATE: 18-MAY-1990  
SEQ ID NO: 5  
LENGTH: 266

Query Match 84.2%; Score 32; DB 6; Length 266;  
Best Local Similarity 66.7%; Pred. No. 3.9e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
DB 52 WLWVIA 6

Db 15 WLWMLA 20

## RESULT 8

5244657-5  
; Patent No. 5244657  
; APPLICANT: KLIEN, MICHEL A.; BOUX, HEATHER A.; COCKLE,  
; STEPHEN A.; LOOSMORE, SHEENA M.; ZEALEY, GAVIN R.  
; TITLE OF INVENTION: GENETIC DETOXIFICATION OF PERTUSSIS  
; TOXIN  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/589,423  
; FILING DATE: 28-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 275,376  
; FILING DATE: 23-NOV-1988  
; SEQ ID NO: 5  
5244657-5  
; LENGTH: 269

## Query Match

Best Local Similarity 84.2%; Score 32; DB 6; Length 269;  
Best Local Similarity 66.7%; Pred. No. 3.9e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLWMLA 6  
| | | | |  
Db 15 WLWMLA 20

## RESULT 9

5433945-5  
; Patent No. 5433945  
; APPLICANT: KLEIN, MICHEL H.; BOUX, HEATHER A.; COCKLE,  
; STEPHEN A.; LOOSMORE, SHEENA M.; ZEALEY, GAVIN R.  
; TITLE OF INVENTION: IMMUNOPROTECTIVE GENETICALLY-DETOXIFIED  
; MUTANTS OF PERTUSSIS TOXIN  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/979,798  
; FILING DATE: 20-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 589,423  
; FILING DATE: 28-SEP-1990  
; APPLICATION NUMBER: 275,376  
; FILING DATE: 23-NOV-1988  
; SEQ ID NO: 5  
5433945-5  
; LENGTH: 269

## Query Match

Best Local Similarity 84.2%; Score 32; DB 6; Length 269;  
Best Local Similarity 66.7%; Pred. No. 3.9e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLWMLA 6  
| | | | |  
Db 15 WLWMLA 20

## RESULT 10

US-09-122-126B-2  
; Sequence 2, Application US/09122126B  
; Patent No. 6451575  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES  
; FILE REFERENCE: DM6909  
; CURRENT APPLICATION NUMBER: US/09/122,126B  
; CURRENT FILING DATE: 1998-07-24  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 837

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-122-126B-2

## Query Match

Best Local Similarity 84.2%; Score 32; DB 4; Length 837;  
Best Local Similarity 80.0%; Pred. No. 1.2e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWML 5  
| | | | |  
Db 33 WLWML 37

## RESULT 11

US-08-469-260A-534  
; Sequence 534, Application US/08469260A  
; Patent No. 6451578  
; GENERAL INFORMATION:  
; APPLICANT: JOHN N. SIMONS  
; APPLICANT: TAMI J. PILOT-MATIAS  
; APPLICANT: GEORGE J. DAWSON  
; APPLICANT: GEORGE G. SCHLAUDER  
; APPLICANT: SURESH M. DESAI  
; APPLICANT: THOMAS P. LEARY  
; APPLICANT: ANTHONY SCOTT MIERHOFF  
; APPLICANT: JAMES C. ERKER  
; APPLICANT: SHERI L. BUTIK  
; APPLICANT: ISA K. MUSHAMMAR  
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS  
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE  
; NUMBER OF SEQUENCES: 716  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D  
; STREET: 100 ABBOTT PARK ROAD  
; CITY: ABBOTT PARK  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064-3500  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,260A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/424,550  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: POREBSKI, PRISCILLA E.  
; REGISTRATION NUMBER: 33,207  
; REFERENCE/DOCKET NUMBER: 5527.PC.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 708-937-6365  
; TELEFAX: 708-938-2623  
; INFORMATION FOR SEQ ID NO: 534:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 50 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-469-260A-534

## Query Match

Best Local Similarity 81.6%; Score 31; DB 4; Length 50;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWMLA 6  
| | | | |  
Db 3 WVVWVS 8

## RESULT 12

US-09-316-080-5  
; Sequence 5, Application US/09316080  
; Patent No. 6180366  
; GENERAL INFORMATION:  
; APPLICANT: John C. Royer  
; APPLICANT: Lynne M. Christlanson  
; APPLICANT: Gregory A. Gambetta  
; APPLICANT: Howard Brody  
; APPLICANT: Suzanne M. Otani  
; TITLE OF INVENTION: Methods For Producing Heterologous  
; TITLE OF INVENTION: Polypeptides In Trichothecene-Deficient Filamentous Fungal  
; FILE REFERENCE: 5563.200-US  
; CURRENT APPLICATION NUMBER: US/09/316,080  
; EARLIER FILING DATE: 1999-05-20  
; EARLIER FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Fusarium  
US-09-316-080-5

Query Match 81.6%; Score 31; DB 4; Length 377;  
Best Local Similarity 60.0%; Pred. No. 7.5e+02;

Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWI 5  
1:11:

DB 220 WWWW 224

## RESULT 13

US-09-316-080-2  
; Sequence 2, Application US/09316080  
; Patent No. 6180366  
; GENERAL INFORMATION:  
; APPLICANT: John C. Royer  
; APPLICANT: Lynne M. Christlanson  
; APPLICANT: Gregory A. Gambetta  
; APPLICANT: Howard Brody  
; APPLICANT: Suzanne M. Otani  
; APPLICANT: Wendy T. Yoder  
; TITLE OF INVENTION: Methods For Producing Heterologous  
; TITLE OF INVENTION: Polypeptides In Trichothecene-Deficient Filamentous Fungal  
; FILE REFERENCE: 5563.200-US  
; CURRENT APPLICATION NUMBER: US/09/316,080  
; CURRENT FILING DATE: 1999-05-20  
; EARLIER FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 380  
; TYPE: PRT  
; ORGANISM: Fusarium  
US-09-316-080-2

Query Match 81.6%; Score 31; DB 4; Length 380;  
Best Local Similarity 60.0%; Pred. No. 7.6e+02;

Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWI 5  
1:11:

DB 220 WWWW 224

## RESULT 14

US-08-952-365-2  
; Sequence 2, Application US/08952365  
; Patent No. 6274311  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: A Method and Nucleotide Sequence for  
; Patent No. 6274311  
; TITLE OF INVENTION: Transforming Microorganisms  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BERSKIN & PARR  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3T2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kurdydyk, Linda M.  
; REGISTRATION NUMBER: 34,971  
; REFERENCE/DOCKET NUMBER: 6580-63  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 364-7311  
; TELEFAX: (416) 361-1398  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 438 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-952-365-2

Query Match 81.6%; Score 31; DB 4; Length 438;  
Best Local Similarity 60.0%; Pred. No. 8.7e+02;

Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWI 5  
1:11:

DB 133 WWWW 137

## RESULT 15

US-08-663-566A-19  
; Sequence 19, Application US/08663566A  
; Patent No. 5853733  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D  
; APPLICANT: Macdonald, Richard D  
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys  
; TITLE OF INVENTION: and Uses Thereof  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/663,566A  
 FILING DATE: June 13, 1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P  
 REGISTRATION NUMBER: 28,678  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212)278-0400  
 TELEFAX: (212)391-0526  
 TELEX: 422523  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 705 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-08-663-566A-19

Query Match 81.6% Score 31: DB 2: Length 705;  
 Best Local Similarity 66.7% Pred No. 1.4e+03;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 WLWIA 6  
 1 1 1 1  
 Db 640 WYVWLA 645

Search completed: January 3, 2003, 15:34:32  
 Job time : 10 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:32:11 : Search time 4.77273 Seconds  
(without alignments)  
23.825 Million cell updates/sec

Title: US-09-543-188a-27

Perfect score: 38

Sequence: 1 WLWVIA 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 117078 seqs, 18951520 residues

Number of hits satisfying chosen parameters: 117078

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PC1\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PC1US\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	97.4	230	9	US-09-738-626-3571
2	34	89.5	49	10	US-09-995-494-75
3	34	89.5	420	10	US-08-815-242-11260
4	33	86.8	296	10	US-09-741-669-378
5	33	86.8	735	10	US-09-815-242-12089
6	32	84.2	75	8	US-08-424-550B-478
7	32	84.2	110	10	US-09-867-550-298
8	32	84.2	560	10	US-09-815-242-11746
9	32	84.2	837	12	US-10-052-586-352
10	31	81.6	50	8	US-08-424-550B-534
11	31	81.6	73	10	US-09-864-761-37732
12	31	81.6	86	10	US-09-864-761-48030
13	31	81.6	96	10	US-09-864-761-33793
14	31	81.6	126	10	US-09-864-761-48848
15	31	81.6	401	10	US-09-871-874-11
16	31	81.6	438	10	US-08-894-993-2
17	31	81.6	441	10	US-09-871-874-21
18	31	81.6	441	10	US-09-895-686-1
19	31	81.6	445	9	US-09-992-598-177

20	31	81.6	445	9	US-09-989-293A-177	Sequence 177, App
21	31	81.6	445	9	US-10-063-547-32	Sequence 32, Appl
22	31	81.6	445	9	US-09-989-735-177	Sequence 177, App
23	31	81.6	445	9	US-09-990-444-177	Sequence 177, App
24	31	81.6	445	9	US-09-989-730-177	Sequence 177, App
25	31	81.6	445	9	US-09-990-436-177	Sequence 177, App
26	31	81.6	445	9	US-09-991-181-177	Sequence 177, App
27	31	81.6	445	9	US-09-993-687-177	Sequence 177, App
28	31	81.6	445	10	US-09-989-722-177	Sequence 177, App
29	31	81.6	445	10	US-09-989-723-177	Sequence 177, App
30	31	81.6	445	10	US-09-989-729-177	Sequence 177, App
31	31	81.6	445	10	US-09-989-727-177	Sequence 177, App
32	31	81.6	445	10	US-09-989-721-177	Sequence 177, App
33	31	81.6	445	10	US-09-989-732-177	Sequence 177, App
34	31	81.6	445	10	US-09-989-732-177	Sequence 177, App
35	31	81.6	445	10	US-09-990-442-177	Sequence 177, App
36	31	81.6	445	10	US-09-991-163-177	Sequence 177, App
37	31	81.6	445	10	US-09-993-604-177	Sequence 177, App
38	31	81.6	445	10	US-09-990-456-177	Sequence 177, App
39	31	81.6	445	10	US-09-989-721-177	Sequence 177, App
40	31	81.6	445	12	US-10-006-867-32	Sequence 32, Appl
41	31	81.6	445	12	US-10-052-586-148	Sequence 148, App
42	31	81.6	446	9	US-09-738-626-6356	Sequence 6356, Ap
43	31	81.6	446	10	US-09-871-874-10	Sequence 10, Appl
44	31	81.6	446	10	US-09-951-780-2	Sequence 2, Appl
45	31	81.6	451	10	US-09-871-874-9	Sequence 9, Appl

## ALIGNMENTS

RESULT 1

US-09-738-626-3571

Sequence 3571, Application US/09738626

Publication No. US20020197605A1

GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI

APPLICANT: ANDO, SEIKO

APPLICANT: HAYASHI, MIKIRO

APPLICANT: OCHIAI, KEIKO

APPLICANT: YOKOI, HARUHIKO

APPLICANT: TATEISHI, NAKO

APPLICANT: SENO, AKIHIRO

APPLICANT: IKEDA, MASATO

APPLICANT: OZAKI, AKIO

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-125

CURRENT APPLICATION NUMBER: US/09/738, 626

CURRENT FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: JP 99/377484

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: JP 00/159162

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: JP 00/280988

PRIOR FILING DATE: 2000-08-03

NUMBER OF SEQ ID NOS: 7059

SOFTWARE: PatentIn ver. 3.0

SEQ ID NO 3571

LENGTH: 230

TYPE: PRT

ORGANISM: Corynebacterium glutamicum

US-09-738-626-3571

Query Match

Best Local Similarity 97.4%; Score 37; DB 9; Length 230;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6

11:1111

DB 170 WLWVIA 175

```

RESULT 2
US-09-995-494-75
; Sequence 75, Application us/09995494
; Patent No. US20020127578A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Caferkey, Robert
; APPLICANT: Ali, Shujath
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; APPLICANT: Chen, Sei-Yu
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0293
; CURRENT APPLICATION NUMBER: US/09/995,494
; CURRENT FILING DATE: 2001-11-27
; PRIOR FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-995-494-75

Query Match
Best Local Similarity 89.5%; Score 34; DB 10; Length 49;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVIA 6
|:|:|
Db 8 WVVWVA 13

RESULT 3
US-09-815-242-11260
; Sequence 11260, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11260
; LENGTH: 420
; TYPE: PRT

```

```

; ORGANISM: Haemophilus influenzae
US-09-815-242-11260

Query Match
Best Local Similarity 89.5%; Score 34; DB 10; Length 420;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVI 5
|:|:|
Db 10 WLWVI 14

RESULT 4
US-09-741-669-378
; Sequence 378, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 378
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-378

Query Match
Best Local Similarity 86.8%; Score 33; DB 10; Length 296;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVI 5
|:|:|
Db 44 WLWVI 48

RESULT 5
US-09-815-242-12089
; Sequence 12089, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12089
; LENGTH: 420
; TYPE: PRT

```

PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 12089  
LENGTH: 735  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-815-242-12089

Query Match 86.8%; Score 33; DB 10; Length 735;  
Best Local Similarity 80.0%; Pred. No. 2.3e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLWVI 5  
Db 469 WLWVI 473

## SUBT 6

US-08-424-550B-478  
Sequence 478, Application US/08424550B  
Patent No. US20020119447A1  
GENERAL INFORMATION:  
APPLICANT: JOHN N. SIMONS  
APPLICANT: TAMU J. PILOT-MATIAS  
APPLICANT: GEORGE J. DANSON  
APPLICANT: GEORGE G. SCHLAUDER  
APPLICANT: SURESH M. DESAI  
APPLICANT: THOMAS P. LEARY  
APPLICANT: ANTHONY SCOTT MUEHROFF  
APPLICANT: JAMES C. ERKER  
APPLICANT: SHERI L. BUTIK  
APPLICANT: ISA K. MUSHAWAR  
TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS  
TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE  
NUMBER OF SEQUENCES: 716  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D  
STREET: 100 ABBOTT PARK ROAD  
CITY: ABBOTT PARK  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/424,550B  
FILING DATE:  
CLASSIFICATION: 435435  
ATTORNEY/AGENT INFORMATION:  
NAME: POREBSKI, PRISCILLA E.  
REGISTRATION NUMBER: 33,207  
REFERENCE/DOCKET NUMBER: 5527.PC.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 708-937-6365  
TELEFAX: 708-938-2623  
INFORMATION FOR SEQ ID NO: 478:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 75 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-424-550B-478

Query Match 84.2%; Score 32; DB 8; Length 75;  
Best Local Similarity 66.7%; Pred. No. 39;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
Db 2 WLWVIA 7

## RESULT 7

US-09-867-550-298  
Sequence 298, Application US/09867550  
Patent No. US20020082206A1  
GENERAL INFORMATION:  
APPLICANT: Leach, Martin D.  
APPLICANT: Mehrihan, Fuad  
APPLICANT: Conley, Pamela  
APPLICANT: Law, Debbie  
APPLICANT: Topper, James  
TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a  
TITLE OF INVENTION: Thereby  
FILE REFERENCE: 21402-013 (Cura-313)  
CURRENT APPLICATION NUMBER: US/09/867,550  
CURRENT FILING DATE: 2001-09-20  
PRIOR APPLICATION NUMBER: USSN 60/208,427  
PRIOR FILING DATE: 2000-05-30  
NUMBER OF SEQ ID NOS: 2125  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 298  
LENGTH: 110  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (37)  
OTHER INFORMATION: wherein Xaa may be any one of Ser or Pro or Thr or Ala  
US-09-867-550-298

Query Match 84.2%; Score 32; DB 10; Length 110;  
Best Local Similarity 66.7%; Pred. No. 55;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLWVIA 6  
Db 84 WLWVIA 89

## RESULT 8

US-09-815-242-11746  
Sequence 11746, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; NUMBER OF SEQ ID NOS: 14110  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 11746  
 ; LENGTH: 560  
 ; TYPE: PRT  
 ; ORGANISM: Klebsiella pneumoniae  
 US-09-815-242-11746

Query Match 84.28; Score 32; DB 10; Length 560;  
 Best Local Similarity 83.38; Pred. No. 2.5e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLYVIA 6  
 |||||  
 Db 146 WLSMIA 151

RESULT 9  
 ; 10-052-586-352  
 ; Sequence 352, Application US/10052586  
 ; Patent No. US20020127584A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Chen, Jian  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3430R1C1  
 ; CURRENT APPLICATION NUMBER: US/10/052,586  
 ; CURRENT FILING DATE: 2002-01-15  
 ; PRIOR APPLICATION NUMBER: 60/059263  
 ; PRIOR FILING DATE: 1997-09-18  
 ; PRIOR APPLICATION NUMBER: 60/059266  
 ; PRIOR FILING DATE: 1997-09-18  
 ; PRIOR APPLICATION NUMBER: 60/062250  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/063120  
 ; PRIOR FILING DATE: 1997-10-24  
 ; PRIOR APPLICATION NUMBER: 60/063121  
 ; PRIOR FILING DATE: 1997-10-24  
 ; PRIOR APPLICATION NUMBER: 60/063486  
 ; PRIOR FILING DATE: 1997-10-21  
 ; PRIOR APPLICATION NUMBER: 60/063540  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/063541  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/063544  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/063564  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/063734  
 ; PRIOR FILING DATE: 1997-10-29  
 ; PRIOR APPLICATION NUMBER: 60/063870  
 ; PRIOR FILING DATE: 1997-10-31  
 ; PRIOR APPLICATION NUMBER: 60/064103  
 ; PRIOR FILING DATE: 1997-10-31  
 ; PRIOR APPLICATION NUMBER: 60/065311  
 ; PRIOR FILING DATE: 1997-11-13  
 ; PRIOR APPLICATION NUMBER: 60/066120  
 ; PRIOR FILING DATE: 1997-11-21  
 ; PRIOR APPLICATION NUMBER: 60/066466  
 ; PRIOR FILING DATE: 1997-11-24  
 ; PRIOR APPLICATION NUMBER: 60/066772

; PRIOR FILING DATE: 1997-11-24  
 ; PRIOR APPLICATION NUMBER: 60/069335  
 ; PRIOR FILING DATE: 1997-12-11  
 ; PRIOR APPLICATION NUMBER: 60/069425  
 ; PRIOR FILING DATE: 1997-12-12  
 ; PRIOR APPLICATION NUMBER: 60/069870  
 ; PRIOR FILING DATE: 1997-12-17  
 ; PRIOR APPLICATION NUMBER: 60/068017  
 ; PRIOR FILING DATE: 1997-12-18  
 ; PRIOR APPLICATION NUMBER: 60/077450  
 ; PRIOR FILING DATE: 1998-03-10  
 ; PRIOR APPLICATION NUMBER: 60/077632  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/077649  
 ; PRIOR FILING DATE: 1998-03-11  
 ; PRIOR APPLICATION NUMBER: 60/078886  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/078939  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/079664  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/079786  
 ; PRIOR FILING DATE: 1998-03-27  
 ; PRIOR APPLICATION NUMBER: 60/080107  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080194  
 ; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080327  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080333  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/081049  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081070  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081195  
 ; PRIOR FILING DATE: 1998-04-09  
 ; PRIOR APPLICATION NUMBER: 60/081838  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/082568  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082569  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082704  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082797  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/083322  
 ; PRIOR FILING DATE: 1998-04-28  
 ; PRIOR APPLICATION NUMBER: 60/083495  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083496  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083499  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083559  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/084366  
 ; PRIOR FILING DATE: 1998-05-05  
 ; PRIOR APPLICATION NUMBER: 60/084414  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084639  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084640  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084643  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15

; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/086023  
 ; PRIOR FILING DATE: 1998-05-18  
 ; PRIOR APPLICATION NUMBER: 60/086392  
 ; PRIOR FILING DATE: 1998-05-22  
 ; PRIOR APPLICATION NUMBER: 60/086486  
 ; PRIOR FILING DATE: 1998-05-22  
 ; PRIOR APPLICATION NUMBER: 60/087098  
 ; PRIOR FILING DATE: 1998-05-28  
 ; PRIOR APPLICATION NUMBER: 60/087208  
 ; PRIOR FILING DATE: 1998-05-28  
 ; PRIOR APPLICATION NUMBER: 60/087609  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087759  
 ; PRIOR FILING DATE: 1998-06-02  
 ; PRIOR APPLICATION NUMBER: 60/087827  
 ; PRIOR FILING DATE: 1998-06-03  
 ; PRIOR APPLICATION NUMBER: 60/088025  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088028  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088029  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088033  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088167  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088202  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088212  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088217  
 ; PRIOR FILING DATE: 1998-06-05  
 ; PRIOR APPLICATION NUMBER: 60/088326  
 ; PRIOR FILING DATE: 1998-06-04  
 ; PRIOR APPLICATION NUMBER: 60/088655  
 ; PRIOR FILING DATE: 1998-06-09  
 ; PRIOR APPLICATION NUMBER: 60/088722  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088738  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088740  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088811  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088824  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088825  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088826  
 ; PRIOR FILING DATE: 1998-06-10  
 ; PRIOR APPLICATION NUMBER: 60/088861  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088863  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/088876  
 ; PRIOR FILING DATE: 1998-06-11  
 ; PRIOR APPLICATION NUMBER: 60/089090  
 ; PRIOR FILING DATE: 1998-06-12  
 ; PRIOR APPLICATION NUMBER: 60/089105  
 ; PRIOR FILING DATE: 1998-06-12  
 ; PRIOR APPLICATION NUMBER: 60/089512  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089514  
 ; PRIOR FILING DATE: 1998-06-16  
 ; PRIOR APPLICATION NUMBER: 60/089538  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089598  
 ; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089653

; PRIOR FILING DATE: 1998-06-17  
 ; PRIOR APPLICATION NUMBER: 60/089908  
 Query Match 84.2%; Score 32; DB 12; Length 837;  
 Best Local Similarity 80.0%; Pred. No. 3.7e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 WLWVI 5  
 Db 33 WLWVL 37  
 RESULT 10  
 US-08-424-550B-534  
 ; Sequence 534, Application US/08424550B  
 ; Patent No. US20020119447A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOHN N. SIMONS  
 ; APPLICANT: TAMU J. PILOT-MATTIAS  
 ; APPLICANT: GEORGE J. DAWSON  
 ; APPLICANT: GEORGE G. SCHLAUDER  
 ; APPLICANT: SURESH M. DESAI  
 ; APPLICANT: THOMAS P. LEARY  
 ; APPLICANT: ANTHONY SCOTT MIERHOFF  
 ; APPLICANT: JAMES C. ERKER  
 ; APPLICANT: SHERI L. BUIK  
 ; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS  
 ; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE  
 ; NUMBER OF SEQUENCES: 716  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D  
 ; STREET: 100 ABBOTT PARK ROAD  
 ; CITY: ABBOTT PARK  
 ; STATE: IL  
 ; COUNTRY: USA  
 ; ZIP: 60064-3500  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/424,550B  
 ; FILING DATE:  
 ; CLASSIFICATION: 435435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: POREBSKI, PRISCILLA E.  
 ; REGISTRATION NUMBER: 33,207  
 ; TELEPHONE: 708-937-6365  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEFAX: 708-938-2623  
 ; INFORMATION FOR SEQ. ID NO: 534:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 50 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-424-550B-534  
 Query Match 81.6%; Score 31; DB 8; Length 50;  
 Best Local Similarity 50.0%; Pred. No. 38;  
 Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 WLWVIA 6  
 Db 3 WVVWVS 8  
 RESULT 11  
 US-09-864-761-37732  
 ; Sequence 37732, Application US/09864761

```

Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 37732
LENGTH: 73
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AI009182.13, SIGNAL = 0.99
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EST_HUMAN HIT: W04502.1, EVALU6 6.00e-02
OTHER INFORMATION: SWISSPROT HIT: P05882, EVALU6 3.90e-01
US-09-864-761-37732

```

Query Match 81.6%; Score 31; DB 10; Length 73;  
 Best Local Similarity 66.7%; Pred. No. 54;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 WLVWIA 6  
 Db 63 WLVWIA 68

```

RESULT 12
US-09-864-761-48030
Sequence 48030, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 48030
LENGTH: 86
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AI009182.12, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
OTHER INFORMATION: EST_HUMAN HIT: BE398021.1, EVALU6 2.00e-03
OTHER INFORMATION: SWISSPROT HIT: P05882, EVALU6 5.20e-01
US-09-864-761-48030

```

Query Match 81.6%; Score 31; DB 10; Length 86;  
 Best Local Similarity 66.7%; Pred. No. 63;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLVWIA 6  
1:111  
Db 76 WLVWIA 81

```
RESULT 13
US-09-864-761-33793
; Sequence 33793, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33793
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009307.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.97
```

```
; OTHER INFORMATION: SWISSPROT HIT: Q23445, EVALUE 3.80e+00
; OTHER INFORMATION: EST HUMAN HIT: A0119728.1, EVALUE 1.00e-38
; OTHER INFORMATION: EST_HUMAN HIT: BE069876.1, EVALUE 8.00e-03
US-09-864-761-33793

Query Match      81.6%; Score 31; DB 10; Length 96;
Best Local Similarity 83.3%; Pred. No. 70;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 WLVWIA 6
1:1111
Db 80 WLVWIA 85

RESULT 14
US-09-864-761-48848
; Sequence 48848, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecmca-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48848
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010368.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
```

OTHER INFORMATION: SWISSPROT HIT: P27958, EVALUE 7.50e-01  
 OTHER INFORMATION: EST\_HUMAN HIT: A1472010.1, EVALUE 2.00e-57  
 US-09-864-761-48848

Query Match 81.6%; Score 31; DB 10; Length 126;  
 Best Local Similarity 80.0%; Pred. No. 90;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWMI 5  
 1:111  
 Db 110 WYVMI 114

RESULT 15  
 US-09-871-874-11  
 : Sequence 11, Application US/09871874  
 : Patent No. US20020081655A1  
 : GENERAL INFORMATION:  
 : APPLICANT: SAVITZKY, Kinneret  
 : APPLICANT: TOPORIK, Amir  
 : APPLICANT: MINTZ, Liat  
 : TITLE OF INVENTION: Splice Variant of mGluR  
 : FILE REFERENCE: 2786-0176P  
 : CURRENT APPLICATION NUMBER: US/09/871, 874  
 : CURRENT FILING DATE: 2001-09-04  
 : NUMBER OF SEQ ID NOS: 21  
 : SOFTWARE: PatentIn Ver. 2.1  
 : SEQ ID NO 11  
 : LENGTH: 401  
 : TYPE: PRT  
 : ORGANISM: Homo sapiens  
 US-09-871-874-11

Query Match 81.6%; Score 31; DB 10; Length 401;  
 Best Local Similarity 80.0%; Pred. No. 2.6e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLWMI 5  
 1:111  
 Db 256 WYVMI 260

Search completed: January 3, 2003, 15:52:07  
 Job time : 4.77273 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:00 ; Search time 9.95455 seconds  
(without alignments)  
57.944 Million cell updates/sec

Title: US-09-543-188A-23

Perfect score: 36

Sequence: 1 IFFWLK 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	91.7	101	2	B87530
2	32	88.9	73	2	D81622
3	32	88.9	73	2	S59082
4	32	88.9	374	2	T32207
5	32	88.9	423	2	E90569
6	32	88.9	544	2	T51162
7	32	88.9	660	2	A63226
8	32	88.9	988	2	H71338
9	31	86.1	86	2	AD0850
10	31	86.1	162	2	D69381
11	31	86.1	247	2	T33469
12	31	86.1	258	2	AG2542
13	31	86.1	262	2	S75891
14	31	86.1	267	2	A72238
15	31	86.1	287	2	T05338
16	31	86.1	388	2	S57526
17	31	86.1	398	2	H70142
18	31	86.1	400	2	H97550
19	31	86.1	400	2	AH2770
20	31	86.1	449	2	C86496
21	31	86.1	449	2	D72127
22	31	86.1	449	2	AD1544
23	31	86.1	474	2	S41117
24	31	86.1	495	1	S25942
25	31	86.1	497	1	S53834
26	31	86.1	500	2	F88921
27	31	86.1	507	1	A30828
28	31	86.1	513	2	B70478
29	31	86.1	532	2	G82872

30	31	86.1	564	2	T20191	hypothetical prote
31	31	86.1	597	2	T27497	hypothetical prote
32	31	86.1	639	1	S45776	uracil transport p
33	31	86.1	669	2	I54205	galactosylceramida
34	31	86.1	738	2	S10659	membrane protein t
35	31	86.1	826	1	A31822	villin - chicken
36	31	86.1	951	1	SYECVT	valine-tRNA ligase
37	31	86.1	951	2	C91283	valine-tRNA synthet
38	31	86.1	951	2	E86124	valine-tRNA synthet
39	31	86.1	951	2	AC1061	valine-tRNA ligase
40	31	86.1	955	2	F84972	valine-tRNA ligase
41	31	86.1	963	2	T12087	H+-exporting ATPase
42	31	86.1	965	2	AE0418	valine-tRNA ligase
43	31	86.1	966	2	T52413	H+-exporting ATPase
44	31	86.1	1101	2	T33153	hypothetical prote
45	31	86.1	1113	2	S48495	probable membrane

## ALIGNMENTS

```
RESULT 1
B87530
hypothetical protein CC2267 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: B87530
R:NIerman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:1125647
A:Accession: B87530
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-101 <STO>
A:Cross-references: GB:AE005673; NID:g13423780; PIDN:AAK24238.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2267

Query Match          91.7%; Score 33; DB 2; Length 101;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWLK 6
Db 6 VFFWLK 11

RESULT 2
D81622
hypothetical protein CP0010 [imported] - Chlamydia pneumoniae (strain AR39)
C:Species: Chlamydia pneumoniae
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C:Accession: D81622
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzbe
r, J.; White, O.; Salzberg, S.L.; Venter, J.C.; Fraser, C.
Proc. Natl. Acad. Sci. U.S.A. 98, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39
A:Reference number: AB1500; MUID:20150255; PMID:10684935
A:Accession: D81622
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-35 <RFA>
A:Cross-references: GB:AE002165; GB:AE002161; NID:g7188948; PIDN:AAF37906.1; PID:g718
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP0010

Query Match          88.9%; Score 32; DB 2; Length 35;
Best Local Similarity 83.3%; Pred. No. 13;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 IFFWIK 6  
11111  
DB 12 IFFWIK 17

## RESULT 3

SS9082  
hypothetical protein 73 - red alga (*Chondrus crispus*) mitochondrion  
C:Species: mitochondrion *Chondrus crispus* (Carraheen)  
C>Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 21-Jul-2000  
C:Accession: SS9082  
R:Leblanc, C.; Boyen, C.; Richard, O.; Bonnard, G.; Grienenberger, J.M.; Kloareg, B.  
J. Mol. Biol. 250, 484-495, 1995  
A>Title: Complete sequence of the mitochondrial DNA of the rhodophyte *Chondrus crispus*  
A:Reference number: SS9078; MUID:95341681; PMID:7616569  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-73 <LEB>  
A:Cross-references: EMBL:Z47547; NID:q1019057; PIDN:CA87598.1; PID:q1334478  
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1995

C:Genetics:  
A:Genome: mitochondrion  
A:Genetic code: SCC3  
C:Keywords: mitochondrion

Query Match 88.9%; Score 32; DB 2; Length 73;  
Best Local Similarity 83.3%; Pred. No. 26;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 IFFWIK 6  
11111  
DB 38 IFFWIK 43

## RESULT 4

T32207  
hypothetical protein T03D3.5 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T32207  
R:Murray, J.; Wohlmann, P.; Bauer, C.; Blewald, T.  
submitted to the EMBL Data Library, September 1997  
A:Description: The sequence of *C. elegans* cosmid T03D3.  
A:Reference number: Z21136  
A:Accession: T32207  
A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-374 <MUR>  
A:Cross-references: EMBL:AF022980; PIDN:AAB69912.1; GSPDB:GN00023; CESP:T03D3.5  
A:Experimental source: strain Bristol N2; clone T03D3  
C:Genetics:  
A:Gene: CESP:T03D3.5  
A:Map position: 5  
A:Introns: 17/3; 28/3; 43/2; 93/3; 155/3; 209/2; 244/2

Query Match 88.9%; Score 32; DB 2; Length 374;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 FFWIK 6  
11111  
DB 36 FFWIK 40

## RESULT 5

E90569  
conserved hypothetical protein MYPV\_4610 [imported] - *Mycoplasma pulmonis* (strain UAB CT  
C:Species: *Mycoplasma pulmonis*.  
C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C:Accession: E90569

R:Chamblaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer,  
Nucleic Acids Res. 29, 2145-2153, 2001  
A>Title: The complete genome sequence of the murine respiratory pathogen *Mycoplasma p*  
A:Reference number: A95512; MUID:21267165; PMID:11353084  
A:Accession: E90569  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-423 <KUR>

A:Cross-references: GB:AL445566; PID:q14089875; PIDN:CAC13634.1; GSPDB:GN00153  
A:Experimental source: strain UAB CT1P  
C:Genetics:  
A:Gene: MYPV\_4610  
A:Genetic code: SCC3

Query Match 88.9%; Score 32; DB 2; Length 423;  
Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWIK 6  
11111  
DB 390 IFFWIK 395

## RESULT 6

T51162  
aminoglycoside acetyltransferase regulator [imported] - *Providencia stuartii*  
C:Species: *Providencia stuartii*  
C>Date: 28-Jul-2000 #sequence\_revision 28-Jul-2000 #text\_change 02-Sep-2000  
C:Accession: T51162  
R:Macinga, D.R.; Cook, G.M.; Poole, R.K.; Rather, P.N.  
J. Bacteriol. 180, 128-135, 1998  
A>Title: Identification and characterization of *aarf*, a locus required for production  
arti.

A:Reference number: Z25318; MUID:98083065; PMID:9422602  
A:Accession: T51162  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-544 <MAC>  
A:Cross-references: EMBL:AF002165; PIDN:AAB96577.1  
A:Experimental source: strain PR50  
C:Genetics:  
A:Gene: *aarf*  
C:Superfamily: Synchocystis ABC transporter slr1919

Query Match 88.9%; Score 32; DB 2; Length 544;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 FFWIK 6  
11111  
DB 41 FFWIK 45

## RESULT 7

AE3226  
conserved hypothetical protein Atus541 [imported] - *Agrobacterium tumefaciens* (strain  
C:Species: *Agrobacterium tumefaciens*  
C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AE3226  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo  
erage, G.; Gillet, W.; Grant, C.; Guentherer, D.; Kutyavin, T.; Levy, R.; Li, M.; MCI  
; Kar, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
ster, E.W.

A>Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AE3226  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-660 <KUR>  
A:Cross-references: GB:AE008687; PIDN:AAL46227.1; PID:q17744003; GSPDB:GN00188  
A:Experimental source: strain C58 (Dupont)

C:Genetics:  
A:Gene: Atu5541  
A:Genome: plasmid

Query Match 88.9%; Score 32; DB 2; Length 660;  
Best Local Similarity 83.3%; Pred. No. 2.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6  
DB 358 IFFWIK 363

## RESULT 8

conserved hypothetical protein TP0325 - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 21-Jan-2000  
C:Accession: H71338

A:Author: C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
Smith, J.; Khaliq, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo  
ney, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:98332770; PMID:9665876

A:Accession: H71338

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-968 <COL>

A:Cross-references: GB:AE001212; GB:AE00520; NID:g3322597; PIDN:AAC65312.1; PID:g332260

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0325

C:Superfamily: syphilis spirochete conserved hypothetical protein TP0325

Query Match 88.9%; Score 32; DB 2; Length 388;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFWIK 6  
DB 747 FFWIK 751

## RESULT 9

AD0850

doubtful CDS found within S. typhi pathogenicity island [imported] - Salmonella enterica  
C:Species: Salmonella enterica subsp. enterica serovar typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001

C:Accession: AD0850

R:Parhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
Th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moutle, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; PMID:11677608

A:Accession: AD0850

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-86 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD05987.1; PID:g16503958; GSPDB:GN00176

C:Genetics:

A:Gene: STY3003

Query Match 86.1%; Score 31; DB 2; Length 86;  
Best Local Similarity 66.7%; Pred. No. 47;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6  
DB 53 LFFWIK 58

RESULT 10  
D69381  
hypothetical protein AF1052 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999

C:Accession: D69381

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod  
., Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes,  
Smith, H.O.; Moese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: D69381

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-162 <RLE>

A:Cross-references: GB:AE001030; GB:AE000782; NID:g2689353; PIDN:AAB90193.1; PID:g264

Query Match 86.1%; Score 31; DB 2; Length 162;  
Best Local Similarity 50.0%; Pred. No. 84;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6  
DB 68 VFFWIK 73

## RESULT 11

T33469

hypothetical protein F43B10.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 04-Mar-2000

C:Accession: T33469

R:Fullon, R.; Hawkins, J.; Rohlfing, T.

submitted to the EMBL Data Library, October 1998

A:Description: The sequence of C. elegans cosmid F43B10.

A:Reference number: Z21351

A:Accession: T33469

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-247 <FUL>

A:Cross-references: EMBL:AF096500; PIDN:AAC67400.1; GSPDB:GN00028; CESP:F43B10.1

A:Experimental source: strain Bristol N2; clone F43B10

C:Genetics:

A:Gene: CESP:F43B10.1

A:Map position: X

A:Introns: 88/2; 103/3; 128/3

C:Superfamily: Caenorhabditis elegans hypothetical protein F43B10.1

Query Match 86.1%; Score 31; DB 2; Length 247;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 5  
DB 85 IFFWIK 89

## RESULT 12

AG2542

hypothetical protein all7626 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC712

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002

C:Accession: AG2542

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kurita, T.; Sasamoto, S.; Watanabe, A.; Irigu  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata  
DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AG2542  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-258 <KUR>  
A:Cross-references: GB:AF003602; PIDN:BAW77269.1; PID:q17134711; GSPDB:GN00181  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all7626  
A:Genome: plasmid

Query Match 86.1%; Score 31; DB 2; Length 258;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWI 5  
|||||  
DB 20 IFFWI 24

ULT 13  
891

hypothetical protein slr1174 - *Synechocystis* sp. (strain PCC 6803)  
C:Species: *Synechocystis* sp.

A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S75891

R:Kaneko, T.; Sato, S.; Kohani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
5.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75891

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-262 <KAN>

A:Cross-references: EMBL:D90913; GB:AB001339; NID:g1653348; PIDN:BAW18350.1; PID:d101908  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match

Best Local Similarity 86.1%; Score 31; DB 2; Length 262;  
100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWI 5  
|||||  
DB 26 IFFWI 30

ULT 14

A72238

hemolysin - *Thermotoga maritima* (strain MSB8)

C:Species: *Thermotoga maritima*

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C:Accession: A72238

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwynn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
C.M.

Nature 399, 323-329, 1999

A:title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: A72238

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-267 <ARN>

A:Cross-references: GB:AE001802; GB:AE000512; NID:g4982133; PIDN:AAW36643.1; PID:g498214

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM1576

C:Superfamily: hemolysin homolog yqxc

Query Match 86.1%; Score 31; DB 2; Length 267;  
Best Local Similarity 80.0%; Pred. No. 1.3e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 2 FFWIK 6  
|||||  
DB 237 FFWIK 241

RESULT 15  
T05338

hypothetical protein F1C12.211 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 23-Jul-1999

C:Accession: T05338

R:Bevan, M.; Terry, N.; Ardiles, W.; Buyschaert, C.; Dasseville, R.; De Clerck, R.;  
ewes, H.W.; Mayer, K.F.X.; Schueller, C.  
submitted to the Protein Sequence Database, April 1998

A:Reference number: Z15408

A:Accession: T05338

A:Molecule type: DNA

A:Residues: 1-287 <BEV>

A:Cross-references: EMBL:AL022224

A:Experimental source: cultivar Columbia; BAC clone F1C12

C:Genetics:

A:Map position: 4

A:Introns: 37/2; 116/2; 243/3; 269/3

A:Note: F1C12.211

Query Match 86.1%; Score 31; DB 2; Length 287;  
Best Local Similarity 80.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 FFWIK 6  
|||||  
DB 53 FFWIK 57

Search completed: January 3, 2003, 15:33:19  
Job time : 10.9545 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:22:10 ; Search time 5.18182 Seconds  
(without alignments)  
48.025 Million call updates/sec

Title: US-09-543-188a-23  
Perfect score: 36  
Sequence: 1 IFFWIK 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112392

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	88.9	544	UBIB_PROST	007443 providencia
2	31	86.1	162	YA52_ARCFU	Q29210 archaeoglob
3	31	86.1	367	NUIC_MESVI	Q9m11 mesostigma
4	31	86.1	473	NORB_PSEST	P98008 pseudomonas
5	31	86.1	495	NUAM_MARPO	P26848 marchantia
6	31	86.1	497	NUAM_MARPO	Q37375 acanthamoeb
7	31	86.1	507	CPT7_RAT	P11715 rattus norv
8	31	86.1	639	FULL_YEAST	P38196 saccharomyc
9	31	86.1	669	GALC_CANFA	P54804 canis famli
10	31	86.1	669	GALC_HUMAN	P54803 homo sapien
11	31	86.1	669	GALC_HUMAN	O02791 macaca mula
12	31	86.1	684	ATY3_HUMAN	Q9h7f0 homo sapien
13	31	86.1	717	TRD1_ECOLI	P09130 escherichia
14	31	86.1	738	TRD2_ECOLI	P22708 escherichia
15	31	86.1	826	VILI_CHICK	P02640 gallus gall
16	31	86.1	951	SVY_ECOLI	P07118 escherichia
17	31	86.1	955	SVY_BUCAT	P57447 buchnera ap
18	31	86.1	1133	WGA2_YEAST	P40578 saccharomyc
19	31	86.1	1256	FLU1_DROME	Q24020 drosophila
20	30	83.3	110	RT10_PEA	P51428 pisum sativ
21	30	83.3	142	RADC_COXBU	O85403 coxiella bu
22	30	83.3	349	YB64_SYNY3	P73771 synechocyst
23	30	83.3	367	NUIC_TOBAC	P06254 nicotiana t
24	30	83.3	368	NUIC_MARPO	P06255 marchantia
25	30	83.3	381	CYB_CARCA	P34665 carcharodon
26	30	83.3	381	CYB_ISUPA	P34671 isurus pauc
27	30	83.3	411	LAG1_YEAST	P38703 saccharomyc
28	30	83.3	470	MM12_HUMAN	P39900 homo sapien
29	30	83.3	537	AREH_SCHPO	O10269 schizosacch
30	30	83.3	642	YAZA_SCHPO	Q09701 schizosacch
31	30	83.3	668	GALC_MOUSE	P54818 mus musculu
32	30	83.3	826	VILI_HUMAN	P09327 homo sapien
33	30	83.3	826	VILI_MOUSE	Q62468 mus musculu

34	30	83.3	865	1	SVY_THEMEA	Q9x2d7 thermotoga
35	30	83.3	875	1	SVY_BORBU	O51680 borrelia bu
36	30	83.3	880	1	SVY_BACBD	Q9x898 bacillus ha
37	30	83.3	880	1	SVY_BACST	P11931 bacillus st
38	30	83.3	880	1	SVY_BACSU	Q05873 bacillus su
39	30	83.3	910	1	SVY_SYNY3	O55522 synechocyst
40	30	83.3	945	1	SVY_NEIMA	Q91x22 neisseria m
41	30	83.3	945	1	SVY_NEIMB	Q91x17 neisseria m
42	30	83.3	953	1	SVY_VIRCH	Q9x673 vibrio chol
43	30	83.3	954	1	SVY_HAEIN	P43834 haemophilus
44	30	83.3	956	1	SVY_TREPA	O83998 treponema p
45	30	83.3	960	1	PMA4_ARATH	Q9su58 arabidopsis

## ALIGNMENTS

RESULT 1  
UBIB\_PROST  
ID UBIB\_PROST STANDARD; PRT; 544 AA.  
AC 007443;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Probable ubiquitinone biosynthesis protein ubib (aminoglycoside  
DE acetyltransferase regulator).  
GN UBIB OR AARF.  
OS Providencia stuartii.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Providencia.  
OX NCBI\_TaxID=588;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PR50;  
RX MEDLINE=98083065; PubMed=9422602;  
RA Macinga D.R., Cook G.M., Poole R.K., Rather P.N.;  
RT Identification and characterization of aarf, a locus required for  
RT production of ubiquinone in Providencia stuartii and Escherichia coli  
RT and for expression of 2'-N-acetyltransferase in P. stuartii.  
RL J. Bacteriol. 180:128-135(1998).  
RN [2]  
RN CHARACTERIZATION.  
RC STRAIN=PR50;  
RX MEDLINE=20416225; PubMed=10960098;  
RA Poon W.W., Davis D.E., Ha H.T., Jonassen T., Rather P.N., Clarke C.F.;  
RT Identification of Escherichia coli ubib, a gene required for the  
RT first monooxygenase step in ubiquinone biosynthesis.  
RL J. Bacteriol. 182:5139-5146(2000).  
CC -1- FUNCTION: REQUIRED, PROBABLY INDIRECTLY, FOR THE HYDROXYLATION OF  
CC 2-OCTAPRENYLPHENOL TO 2-OCTAPRENYL-6-HYDROXY-PHENOL, THE FOURTH  
CC STEP IN UBIQUINONE BIOSYNTHESIS. REQUIRED FOR THE EXPRESSION OF  
CC 2'-N-ACETYLTRANSFERASE.  
CC -1- SIMILARITY: BELONGS TO THE ABC1 FAMILY. UBIB SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AF002165; AAB96577.1; -;  
CC InterPro: IPR004147; ABC1.  
CC DR Pfam: PF03109; ABC1; 1.  
CC Ubiqunone biosynthesis.  
SQ SEQUENCE 544 AA: 62598 MW: E789A4E8185B4F96 CRC64:  
Query Match 88.9%; Score 32; DB 1; Length 544;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 FFWIK 6

Db 41 FFWIK 45

## RESULT 2

YAS2\_ARCFU STANDARD: PRT: 162 AA.  
ID YAS2\_ARCFU  
AC 029210;  
DT 16-OCT-2001 (Rel. 40, Created)  
RT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
GN Hypothetical protein AF1052 precursor.  
OS Archaeoglobus fulgidus.  
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
OX NCBI\_TaxID=2234;  
RN [1]  
RP SEQUENCE FROM N.A.  
STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
MEDLINE=98049343; PubMed=9389475;  
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
Richardson D.L., Kierlavage A.R., Gisham D.E., Kyriakides N.C.,  
Rieschmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,  
Colton M.D., Spriggs T., Arlanch P., Kaine B.P., Sykes S.M.,  
Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
Mason T.M., Olsen G.V., Fraser C.M., Smith H.O., Woese C.R.,  
Venter J.C.;

RT "The complete genome sequence of the hyperthermophilic, sulphate-  
reducing archaeon Archaeoglobus fulgidus.";  
RL Nature 390:364-370(1997).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AE001030; AAB90193.1; -  
DR TIGR: AF1052; -  
KW Hypothetical protein; Signal; Complete proteome.  
RN SIGNAL 1 34  
CHAIN 35 162  
SEQUENCE 162 AA: 18043 MW; A6547DC1139B72B3 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 162;  
Best Local Similarity 50.0%; Pred. No. 39;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWIK 6  
Db 68 YFFWIK 73

## RESULT 3

NOIC\_MESVI STANDARD: PRT: 367 AA.  
ID NOIC\_MESVI  
AC 09MU11;  
DT 16-OCT-2001 (Rel. 40, Created)  
RT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
GN NADH-plastoquinone oxidoreductase chain 1, chloroplast (EC 1.6.5.3).  
OS Mesostigma viride.  
OC Chloroplast.  
CC Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;  
OC Mesostigmatales; Mesostigmataceae; Mesostigma.

OX NCBI\_TaxID=41882;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NIES-296;  
RX MEDLINE=20150907; PubMed=10688199;  
RA Lemieux C., Otis C., Turmel M.;  
RT "Ancestral chloroplast genome in Mesostigma viride reveals an early  
branch of green plant evolution.";  
RL Nature 403:649-653(2000).  
CC -1- CATALYTIC ACTIVITY: NADH + plastoquinone -> NAD(+) + plastoquinol.  
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF166114; AAF43887.1; -  
DR InterPro: IPR001694; Resp\_NADH\_dhl.  
DR Pfam: PF00146; NADHdh; 1.  
DR PROSITE: PS00667; COMPLEX1\_ND1\_1; FALSE\_NEG.  
DR PROSITE: PS00668; COMPLEX1\_ND1\_2; 1.  
KW Oxidoreductase; NAD; Plastoquinone; Chloroplast; Transmembrane.  
SQ SEQUENCE 367 AA: 40369 MW; 158CF9E4A458EB98 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 367;  
Best Local Similarity 100.0%; Pred. No. 85;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWIK 5  
Db 132 IFFWIK 136

## RESULT 4

NORB\_PSES7 STANDARD: PRT: 473 AA.  
ID NORB\_PSES7  
AC P98008;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Nitric-oxide reductase subunit B (EC 1.7.99.7) (Nitric oxide reductase  
DE cytochrome b subunit) (NOR large subunit).  
GN NORB.

OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=316;  
RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.  
RC STRAIN=ATCC 14405 / Zobell;  
RX MEDLINE=94139726; PubMed=7508388;  
RA Zumft W.G., Braun C., Cuypers H.;  
RT "Nitric oxide reductase from Pseudomonas stutzeri. Primary structure  
RT and gene organization of a novel bacterial cytochrome bc complex.";  
RL Eur. J. Biochem. 219:481-490(1994).

RN [2]  
RP EPR SPECTROSCOPY.  
RC STRAIN=ATCC 14405 / Zobell;  
RX MEDLINE=89255095; PubMed=2542222;  
RA Heiss B., Frunzke K., Zumft W.G.;  
RT "Formation of the N-N bond from nitric oxide by a membrane-bound  
RT cytochrome bc complex of nitrate-respiring (denitrifying) Pseudomonas  
RT stutzeri.";  
RL J. Bacteriol. 171:3288-3297(1989).

RN [3]  
RP EPR SPECTROSCOPY.  
RC STRAIN=ATCC 14405 / Zobell;  
RX MEDLINE=98191362; PubMed=9521721;  
RA Cheesman M.R., Zumft W.G., Thomson A.J.;

RT "The MCD and EPR of the heme centers of nitric oxide reductase  
 RT from *Pseudomonas stutzeri*: evidence that the enzyme is  
 RT structurally related to the heme-copper oxidases.";   
 RL Biochemistry 37:3994-4000(1998).  
 CC -1- FUNCTION: COMPONENT OF THE ANAEROBIC RESPIRATORY CHAIN THAT  
 CC TRANSFORMS NITRATE TO DINITROGEN (DENITRIFICATION). NORB IS THE  
 CC CATALYTIC SUBUNIT OF THE ENZYME COMPLEX. SHOWS PROTON PUMP  
 CC ACTIVITY ACROSS THE MEMBRANE IN DENITRIFYING BACTERIAL CELLS. THE  
 CC MONONITROGEN REDUCTION IS PROBABLY COUPLED TO ELECTRON TRANSPORT  
 CC PHOSPHORYLATION  
 CC -1- CATALYTIC ACTIVITY: Nitrous oxide + acceptor + H(2)O = 2 nitric  
 CC oxide + reduced acceptor.  
 CC -1- SUBUNIT: HETERODIMER OF CYTOCHROMES B (LARGE SUBUNIT) AND C (SMALL  
 CC SUBUNIT).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- INDUCTION: BY NITRIC OXIDE (PROBABLE).  
 CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z28384; CAAB2229.1; -  
 DR InterPro: IPR000883; COX1.  
 DR Pfam: PF00115; COX1; 1.  
 DR PROSITE: PS00077; COX1; 1.  
 KW Oxidoreductase; Heme; Iron; Transmembrane; Respiratory chain.  
 FT INIT\_MET 0 0  
 FT METAL 59 59 IRON (LOW-SPIN HEME) (PROBABLE).  
 FT METAL 206 206 IRON B (PROBABLE).  
 FT METAL 257 257 IRON B (PROBABLE).  
 FT METAL 258 258 IRON B (PROBABLE).  
 FT METAL 346 346 IRON (HIGH-SPIN HEME) (PROBABLE).  
 FT METAL 348 348 IRON (LOW-SPIN HEME) (PROBABLE).  
 FT TRANSMEM 18 38 POTENTIAL.  
 FT TRANSMEM 60 80 POTENTIAL.  
 FT TRANSMEM 94 114 POTENTIAL.  
 FT TRANSMEM 144 164 POTENTIAL.  
 FT TRANSMEM 168 188 POTENTIAL.  
 FT TRANSMEM 206 226 POTENTIAL.  
 FT TRANSMEM 242 262 POTENTIAL.  
 FT TRANSMEM 269 289 POTENTIAL.  
 FT TRANSMEM 307 327 POTENTIAL.  
 FT TRANSMEM 347 367 POTENTIAL.  
 FT TRANSMEM 389 409 POTENTIAL.  
 FT TRANSMEM 432 452 POTENTIAL.  
 SQ SEQUENCE 473 AA; 53006 MW; 39E999DC1DD7F757 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 473;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWIK 6  
 DB 434 IFFWIK 439

RESULT 5  
 NU4M\_MARPO STANDARD: PRT; 495 AA.  
 AC P26848;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).  
 GN NDA OR NAD4.  
 OS Marchantia polymorpha (Liverwort).  
 OG Mitochondrion.  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;

OC Marchantiopsida; Marchantiidae; Marchantiales; Marchantiaceae;  
 OC Marchantiaceae; Marchantia.  
 OX NCBI\_TaxID=3197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92114051; PubMed=1731062;  
 RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,  
 RA Akashi K., Kaneage T., Ogura Y., Kohchi T., Ohyama K.;  
 RT "Gene organization deduced from the complete sequence of liverwort  
 RT Marchantia polymorpha mitochondrial DNA. A primitive form of plant  
 RT mitochondrial genome.";   
 RL J. Mol. Biol. 223:1-7(1992).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93247547; PubMed=8483448;  
 RA Nozato N., Oda K., Yamato K., Ohta E., Takemura M., Akashi K.,  
 RA Fukuzawa H., Ohyama K.;  
 RT "Clonally transcribed expression of mitochondrial genes for subunits of  
 RT NADH dehydrogenase, nad5, nad4, nad2, in Marchantia polymorpha.";   
 RL Mol. Genet. 237:343-350(1993).  
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M68929; AAC09398.1; -  
 DR PIR: S25942; S25942.  
 DR InterPro: IPR003918; NADHUB\_OXRED4.  
 DR InterPro: IPR001750; Oxidored\_q1.  
 DR Pfam: PF00361; Oxidored\_q1; 1.  
 DR PRINTS: PR01437; NUOXDRPTASE4.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.  
 SQ SEQUENCE 495 AA; 56311 MW; 0F75894DC6CAAAED4 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 495;  
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWIK 6  
 DB 53 IFFWIK 58

RESULT 6  
 NU4M\_ACACA STANDARD: PRT; 497 AA.  
 AC Q37375;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).  
 GN NDA OR NAD4.  
 OS Acanthamoeba castellanii (Amoeba).  
 OG Mitochondrion.  
 OC Eukaryota; Acanthamoebidae; Acanthamoeba.  
 OX NCBI\_TaxID=5755;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 30010 / NEFF;  
 RX MEDLINE=95147275; PubMed=7844823;  
 RA Burger G., Plante I., Loneragan K.M., Gray M.W.;  
 RT "The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba  
 RT castellanii: complete sequence, gene content and genome  
 RT organization.";   
 RL J. Mol. Biol. 245:522-537(1995).  
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U12386; AAD11826.1; -  
CC InterPro: IPR003918; NADhub\_oxred4.  
CC InterPro: IPR001750; Oxidored\_q1.  
CC Pfam: PF00361; oxidored\_q1.1.  
CC PRINTS: PR01437; NUOXDRDTASE4.  
CC OXIDOREDUCTASE; NAD; Ubiquinone; Mitochondrion.  
SQ SEQUENCE 497 AA; 57768 MW; 9C0B637E65C0F3B5 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 497;  
Best Local Similarity 100.0%; Pred. NO. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 IFFWI 5  
53 IFFWI 57

RESULT 7  
CPT7\_RAT STANDARD: PRT; 507 AA.  
ID CPT7\_RAT  
AC P11715;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Cytochrome P450 17 (EC 1.14.99.9) (CYPXVII) (P450-C17) (Steroid  
DE 17-alpha-hydroxylase/17,20 lyase).  
CN CYP17  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89295447; PubMed=2786990;  
RA Revold H.R., Lorence W.C., McCarthy J.L., Trant J.M., Kagimoto M.,  
RA Waterman M.R., Mason J.I.;  
RT "Rat P450(C17 alpha) from testis: characterization of a full-length  
RT cDNA encoding a unique steroid hydroxylase capable of catalyzing both  
RT delta 4- and delta 5-steroid-17,20-lyase reactions.";  
RL Mol. Endocrinol. 3:968-975(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Testis;  
RA MEDLINE=89076306; PubMed=3264499;  
RA Namiki M., Kitamura M., Buczko E., Dufau M.L.;  
RT "Rat testis P-450(17)alpha cDNA: the deduced amino acid sequence,  
RT expression and secondary structural configuration.";  
RL Blochem. Biophys. Res. Commun. 157:705-712(1988).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley;  
RX MEDLINE=95217329; PubMed=7702752;  
RA Glvens C.R., Zhang P., Balr S.R., Mellon S.H.;  
RT "Transcriptional regulation of rat cytochrome P450c17 expression in  
RT mouse Leydig MA-10 and adrenal Y-1 cells: identification of a single  
RT protein that mediates both basal and CAMP-induced activities.";  
RL DNA Cell Biol. 13:1087-1098(1994).  
RN [4]  
RP SEQUENCE OF 271-507 FROM N.A.  
RX MEDLINE=88280759; PubMed=3260774;  
RA Nishihara M., Winters C.A., Buzko E., Waterman M.R., Dufau M.L.;  
RT "Hormonal regulation of rat Leydig cell cytochrome P-45017 alpha  
RT levels and characterization of a partial length rat P-45017 alpha  
RT cDNA.";  
RL Blochem. Biophys. Res. Commun. 154:151-158(1988).  
RX [5]

RP SEQUENCE OF 273-507 FROM N.A.  
RX MEDLINE=90046678; PubMed=2554289;  
RA Mellon S.H., Vaisse C.;  
RT "cAMP regulates P450csc gene expression by a  
RT cyclonxide-insensitive mechanism in cultured mouse Leydig MA-10  
RT cells.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:7775-7779(1989).  
CC -1- FUNCTION: CONVERSION OF PREGNENOLONE AND PROGESTERONE TO THEIR  
CC 17-ALPHA-HYDROXYLATED PRODUCTS AND SUBSEQUENTLY TO  
CC DEHYDROEPIANDROSTERONE (DHEA) AND ANDROSTERONE. CATALYZES  
CC BOTH THE 17-ALPHA-HYDROXYLATION AND THE 17,20-LYASE REACTION.  
CC INVOLVED IN SEXUAL DEVELOPMENT DURING FETAL LIFE AND AT PUBERTY.  
CC hydroxysteroid + A + H(2O).  
CC -1- ENZYME REGULATION: REGULATED PREDOMINANTLY BY INTRACELLULAR CAMP  
CC LEVELS.  
CC -1- PATHWAY: KEY ENZYME IN STEROIDGENIC PATHWAY.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X14086; CAA32248.1; -  
CC EMBL: X69816; CAA49470.1; -  
CC EMBL: M31681; AAA41777.1; -  
CC EMBL: M22204; AAA41783.1; -  
CC EMBL: M21208; AAA41050.1; -  
CC EMBL: M27282; AAA41779.1; -  
CC PIR: A27659; A27659.  
CC PIR: A30828; A30828.  
CC PIR: S16719; S16719.  
CC HSSP: P00179; 1D76.  
CC InterPro: IPR001128; Cytochrome\_p450.  
CC Pfam: PF00067; P450; 1.  
CC PRINTS: PR00385; P450.  
CC PROSITE: PS00086; CYTOCHROME\_P450; 1.  
KW Electron transport; Oxidoreductase; Monooxygenase; Membrane;  
KW Heme; Steroidogenesis.  
FT BINDING 441 441 HEME (BY SIMILARITY).  
FT CONFLICT 505 506 VS -> LT (IN REF. 4).  
SQ SEQUENCE 507 AA; 57250 MW; A535600F7E6A3999 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 507;  
Best Local Similarity 80.0%; Pred. NO. 1.1e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

2 FFWIK 6  
15 FFWIK 19

Db 15 FFWIK 19

RESULT 8  
FUT11\_YEAST STANDARD: PRT; 639 AA.  
ID FUT11\_YEAST  
AC P38196;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Uridine permease.  
GN FUT11 OR YBL042C OR YBL0406.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=95176707; PubMed=7871888;



RA de Wergifosse P., Jacques B., Jonniaux J.-L., Purnelle B., Skala J.,  
RA Goffeau A.;  
RT "The sequence of a 22.4 kb DNA fragment from the left arm of yeast  
RT chromosome II reveals homologues to bacterial proline synthetase and  
RT murine alpha-actinin, as well as a new permease and a DNA-binding  
RT protein.";  
RL Yeast 10:1489-1496(1994).  
RN  
RP CHARACTERIZATION.  
RX MEDLINE=98146544; PubMed=9485596;  
RA Wagner R., de Montigny J., de Wergifosse P., Souciet J.-L., Potier S.;  
RT "The ORF YBL042 of *Saccharomyces cerevisiae* encodes a uridine  
RT permease.";  
RL FEMS Microbiol. Lett. 159:69-75(1998).  
RN  
RP  
CC CHARACTERIZATION.  
RX MEDLINE=20408929; PubMed=10827169;  
RA Vickers M.F., Yao S.Y., Baldwin S.A., Young J.D., Cass C.E.;  
RT "Nucleoside transporter proteins of *Saccharomyces cerevisiae*.  
RT Demonstration of a transporter (FUT1) with high uridine selectivity  
RT in plasma membranes and a transporter (FUT2) with broad nucleoside  
RT selectivity in intracellular membranes.";  
RL J. Biol. Chem. 275:25931-25938(2000).  
CC  
CC -1- FUNCTION: HIGH-AFFINITY TRANSPORT OF URIDINE.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO THE ATLANTOIN PERMEASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
DR EMBL; X78214; CAAS5059.1; -;  
DR EMBL; Z35803; CAAS4862.1; -;  
DR SGD; S0000138; FUT1.  
DR InterPro: IPR001248; Cyt\_pur\_permease.  
DR Pfam; PF02133; transp\_cyt\_pur; 1.  
DR TIGRfams; TIGR00800; ncs1; 1.  
KW Transmembrane; Transport.  
FT TRANSMEM 163 180 POTENTIAL.  
FT TRANSMEM 201 225 POTENTIAL.  
FT TRANSMEM 260 276 POTENTIAL.  
FT TRANSMEM 284 305 POTENTIAL.  
FT TRANSMEM 368 392 POTENTIAL.  
FT TRANSMEM 417 435 POTENTIAL.  
FT TRANSMEM 461 477 POTENTIAL.  
FT TRANSMEM 484 507 POTENTIAL.  
FT TRANSMEM 538 562 POTENTIAL.  
FT TRANSMEM 573 590 POTENTIAL.  
SQ SEQUENCE 639 AA; 72165 MW; 757325DEF4100399 CRC64;  
Query Match 86.1%; Score 31; DB 1; Length 639;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE (Galactocerebrosidase).  
GN GALC.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxId=9615;  
RN  
RP  
RX MEDLINE=96299640; PubMed=8661004;  
RA Victoria T., Rafi M.A., Wenger D.A.;  
RT "Cloning of the canine GALC cDNA and identification of the mutation  
RT causing globoid cell leukodystrophy in West Highland White and Cairn  
RT terriers.";  
RL Genomics 33:457-462(1996).  
CC  
CC -1- FUNCTION: HYDROLYSES THE GALACTOSE ESTER BONDS OF  
CC GALACTOSYLKERAMIDE, GALACTOSYLSPHINGOSINE, LACTOSYLKERAMIDE, AND  
CC MONOGALACTOSYLGLYCERIDE. ENZYME WITH VERY LOW ACTIVITY  
CC RESPONSIBLE FOR THE LYSOSOMAL CATABOLISM OF GALACTOSYLKERAMIDE, A  
CC MAJOR LIPID IN MYELIN, KIDNEY AND EPITHELIAL CELLS OF SMALL  
CC INTESTINE AND COLON (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: D-galactosyl-N-acylsphingosine + H(2)O = D-  
CC galactose + N-acylsphingosine.  
CC -1- SUBCELLULAR LOCATION: Lysosomal (By similarity).  
CC -1- DISEASE: DEFECTS IN GALC ARE THE CAUSE OF GLOBOID CELL  
CC LEUKODYSTROPHY (GLD). THIS DEFICIENCY RESULTS IN THE INSUFFICIENT  
CC CATABOLISM OF SEVERAL GALACTOLIPIDS THAT ARE IMPORTANT IN THE  
CC PRODUCTION OF NORMAL MYELIN.  
CC -1- SIMILARITY: BELONGS TO FAMILY 59 OF GLYCOSYL HYDROLASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
DR EMBL; L76184; AAB37752.1; -;  
DR InterPro: IPR001286; GH\_59.  
DR Pfam; PF02057; Glyco\_Hydro\_59; 1.  
DR PRINTS; PR00850; GLHYDRLASE59.  
KW Hydrolyase; Glycosidase; Glycoprotein; Signal; Disease mutation.  
FT SIGNAL 1 26  
FT CHAIN 27 669  
FT CARBOHYD 363 363  
FT CARBOHYD 367 367  
FT CARBOHYD 543 543  
FT CARBOHYD 586 586  
FT VARIAT 158 158  
FT VARIANT 639 639  
SQ SEQUENCE 669 AA; 75316 MW; 60E298B024EE154C CRC64;  
Query Match 86.1%; Score 31; DB 1; Length 669;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5  
DB 579 IFFWI 583

RESULT 10  
GALC\_HUMAN  
ID GALC\_HUMAN STANDARD; PRT; 669 AA.  
AC P54803;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Galactocerebrosidase precursor (EC 3.2.1.46) (GALCPRASE)  
DE (Galactosylceramidase) (Galactosylceramide beta-galactosidase)  
GN GALC.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Placenta, and SKIN fibroblast;  
RA MEDLINE=9412808; PubMed=8297359;  
RX Sakai N., Inui K., Fujii N., Fukushima H., Nishimoto J.,  
RA Yanagihara I., Isegawa Y., Iwamatsu A., Okada S.;  
RT "Krabbe disease: isolation and characterization of a full-length cDNA  
for human galactocerebrosidase.";  
RL Biochem. Biophys. Res. Commun. 198;485-491(1994).  
RN [2]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-59 AND 436-454.  
RC TISSUE=Brain, and Testis;  
RA MEDLINE=94108435; PubMed=8281145;  
RX Chen Y.O., Rafi M.A., de Galla G., Wenger D.A.;  
RT "Cloning and expression of cDNA encoding human galactocerebrosidase,  
the enzyme deficient in globoid cell leukodystrophy.";  
Hum. Mol. Genet. 2:1841-1845(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95324938; PubMed=7601472;  
RT Luzi P., Rafi M.A., Wenger D.A.;  
RT "Structure and organization of the human galactocerebrosidase (GALC)  
gene.";  
RT Genomics 26:407-409(1995).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98094242; PubMed=9434153;  
RA Sakai N., Fukushima H., Inui K., Fu L., Nishigaki T., Yanagihara I.,  
RA Tatsumi N., Ozono K., Okada S.;  
RT "Human galactocerebrosidase gene: promoter analysis of the 5'-flanking  
region and structural organization.";  
RL Biochim. Biophys. Acta 1395:62-67(1998).  
RN [5]  
RP SEQUENCE OF 27-45 AND 436-454, AND CHARACTERIZATION.  
RC TISSUE=Urine;  
RX MEDLINE=94002192; PubMed=8399327;  
RA Chen Y.O., Wenger D.A.;  
RT "Galactocerebrosidase from human urine: purification and partial  
characterization.";  
RL Biochim. Biophys. Acta 1170:53-61(1993).  
RN [6]  
RP REVIEW ON GLD MUTATIONS.  
RX MEDLINE=97478285; PubMed=9338580;  
RA Wenger D.A., Rafi M.A., Luzi P.;  
RT "Molecular genetics of Krabbe disease (globoid cell leukodystrophy):  
diagnostic and clinical implications.";  
Hum. Mutat. 10:268-279(1997).  
RN [7]  
RP VARIANTS GLD ALA-302 AND GLY-550.  
RX MEDLINE=96121583; PubMed=8595408;  
RA Tatsumi N., Inui K., Sakai N., Fukushima H., Nishimoto J.,  
RA Yanagihara I., Nishigaki T., Tsukamoto H., Fu L., Tanike M.,  
RA Okada S.;  
RT "Molecular defects in Krabbe disease.";  
Hum. Mol. Genet. 4:1865-1868(1995).  
RN [8]  
RP VARIANTS GLD ASN-528 AND SER-583.  
RX MEDLINE=96198195; PubMed=8786069;  
RA Rafi M.A., Luzi P., Zlotogora J., Wenger D.A.;  
RT "Two different mutations are responsible for Krabbe disease in the  
Druse and Moslem Arab populations in Israel.";  
Hum. Genet. 97:304-308(1996).  
RN [9]  
RP FUNCTION: HYDROLYSES THE GALACTOSE ESTER BONDS OF  
GALACTOSYLCERAMIDE, GALACTOSYLSPHINGOSINE, LACTOSYLCERAMIDE, AND  
MONOGALACTOSYLGLYCERIDE. ENZYME WITH VERY LOW ACTIVITY  
RESPONSIBLE FOR THE LYSOSOMAL CATABOLISM OF GALACTOSYLCERAMIDE, A  
MAJOR LIPID IN MYELIN, KIDNEY AND EPITHELIAL CELLS OF SMALL  
INTESTINE AND COLON. HAS AN OPTIMAL PH BETWEEN 4.0 AND 4.4.  
CC ACTIVITY IS LOST WHEN HEATED AT 52 DEGREES CELSIUS FOR FIVE  
CC MINUTES.

CC -1- CATALYTIC ACTIVITY: D-galactosyl-N-acylsphingosine + H(2)O = D-  
CC galactose + N-acylsphingosine.  
CC -1- SUBCELLULAR LOCATION: Lysosomal.  
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS ARE PRODUCED BY ALTERNATIVE  
CC SPLICING. THE SEQUENCE SHOWN HERE IS THAT OF THE LONG FORM.  
CC -1- TISSUE SPECIFICITY: HIGHEST LEVEL OF ACTIVITY IN TESTES COMPARED  
CC TO BRAIN, KIDNEY, PLACENTA AND LIVER. CAN ALSO BE FOUND IN URINE.  
CC -1- DISEASE: DEFECTS IN GALC ARE THE CAUSE OF GLOBOID CELL  
CC LEUKODYSTROPHY (GLD) (OR KRABBE DISEASE). THIS AUTOSOMAL RECESSIVE  
CC DISORDER DEFICIENCY RESULTS IN THE INSUFFICIENT CATABOLISM OF  
CC SEVERAL GALACTOLIPIDS THAT ARE IMPORTANT IN THE PRODUCTION OF  
CC NORMAL MYELIN. CLINICALLY THE MOST FREQUENT FORM IS THE INFANTILE  
CC FORM. MOST PATIENTS (90%) PRESENT BEFORE SIX MONTHS OF AGE WITH  
CC IRRITABILITY, SPASTICITY, ARREST OF MOTOR AND MENTAL DEVELOPMENT,  
CC AND BOUTS OF TEMPERATURE ELEVATION WITHOUT INFECTION. THIS IS  
CC FOLLOWED BY MYOCLONIC JERKS OF ARMS AND LEGS, OPHTHOMOUS,  
CC HYPERTRONIC FITS, AND MENTAL REGRESSION, WHICH PROGRESSES TO A  
CC SEVERE DECELERATE CONDITION WITH NO VOLUNTARY MOVEMENTS AND DEATH  
CC FROM RESPIRATORY INFECTIONS OR CEREBRAL HYPERTENSIA BEFORE 2  
CC YEARS OF AGE. HOWEVER, A SIGNIFICANT NUMBER OF CASES WITH LATER  
CC ONSET, PRESENTING WITH UNEXPLAINED BLINDNESS, WEAKNESS AND/OR  
CC PROGRESSIVE MOTOR, AND SENSORY NEUROPATHY THAT CAN PROGRESS TO  
CC SEVERE MENTAL INCAPACITY AND DEATH, HAVE BEEN IDENTIFIED.  
CC -1- SIMILARITY: BELONGS TO FAMILY 59 OF GLYCOSYL HYDROLASES.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D25283; BAA04971.1; -;  
DR EMBL; D25284; BAA04972.1; -;  
DR EMBL; L23116; AAA16645.1; -;  
DR EMBL; L38559; AAA80975.1; -;  
DR EMBL; L38544; AAA80975.1; JOINED.  
DR EMBL; L38545; AAA80975.1; JOINED.  
DR EMBL; L38546; AAA80975.1; JOINED.  
DR EMBL; L38547; AAA80975.1; JOINED.  
DR EMBL; L38548; AAA80975.1; JOINED.  
DR EMBL; L38549; AAA80975.1; JOINED.  
DR EMBL; L38550; AAA80975.1; JOINED.  
DR EMBL; L38551; AAA80975.1; JOINED.  
DR EMBL; L38552; AAA80975.1; JOINED.  
DR EMBL; L38553; AAA80975.1; JOINED.  
DR EMBL; L38555; AAA80975.1; JOINED.  
DR EMBL; L38556; AAA80975.1; JOINED.  
DR EMBL; L38557; AAA80975.1; JOINED.  
DR EMBL; L38558; AAA80975.1; JOINED.  
DR EMBL; D84281; BAA24902.1; -;  
DR EMBL; D84282; BAA24902.1; JOINED.  
DR EMBL; D84283; BAA24902.1; JOINED.  
DR EMBL; D84284; BAA24902.1; JOINED.  
DR EMBL; D84266; BAA24902.1; JOINED.  
DR EMBL; D84267; BAA24902.1; JOINED.  
DR EMBL; D84268; BAA24902.1; JOINED.  
DR EMBL; D84269; BAA24902.1; JOINED.  
DR EMBL; D84363; BAA24902.1; JOINED.  
DR EMBL; D84364; BAA24902.1; JOINED.  
DR EMBL; D84365; BAA24902.1; JOINED.  
DR EMBL; D84366; BAA24902.1; JOINED.  
DR EMBL; D84367; BAA24902.1; JOINED.  
DR EMBL; D84368; BAA24902.1; JOINED.  
DR EMBL; D84369; BAA24902.1; JOINED.  
DR EMBL; D84393; BAA24902.1; JOINED.  
DR Genew; HGNC:4115; GALC.  
DR MIM: 606890; -;  
DR MIM: 245200; -;  
DR InterPro: IPR001286; GH\_59.  
DR Pfam: PF02057; Glyco\_hydro\_59; 1.  
DR PRINTS: PR00850; GLHYDRASE59.



16-OCT-2001 (Rel. 40, Created)  
 15-OCT-2001 (Rel. 40, Last sequence update)  
 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Probable cation-transporting ATPase 3 (EC 3.6.3.-) (Fragment).  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Coronary artery;  
 RA Kawabata A., Hiki J. T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,  
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;  
 "NEO human cDNA sequencing project."  
 RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
 RL -1- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2  
 ATPASES). SUBFAMILY V.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

DR EMBL: AK024639; BAB1942.1; ALT\_INIT.  
 DR InterPro: IPR001757; ATPase\_E1-E2.  
 DR Pfam: PF00702; Hydrolase.1.  
 DR PRINTS: PR00119; CATATPASE.  
 DR PROSITE: PS00154; ATPase\_E1\_E2. 1.  
 KM Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.  
 FT NON-TER 1 1  
 FT TRANSMEM 10 29 POTENTIAL.  
 FT TRANSMEM 44 64 POTENTIAL.  
 FT TRANSMEM 537 556 POTENTIAL.  
 FT TRANSMEM 568 585 POTENTIAL.  
 FT TRANSMEM 602 622 POTENTIAL.  
 FT TRANSMEM 659 671 POTENTIAL.  
 FT MOD\_RES 97 97 PHOSPHORYLATION (BY SIMILARITY).  
 FT METAL 482 482 MAGNESIUM (BY SIMILARITY).  
 FT METAL 486 486 MAGNESIUM (BY SIMILARITY).  
 SQ SEQUENCE 684 AA; 76081 MW; D655E494DCEFAE1E7 CRC64;

Query Match 86.1%; Score 31; DB 1; Length 684;  
 Last Local Similarity 80.0%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 FFMIK 6  
 DB 624 FFVWK 628

RESULT 13  
 TRD1\_ECOLI  
 ID TRD1\_ECOLI STANDARD; PRT: 717 AA.  
 AC P09130;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Trd protein.  
 GN TRAD.  
 OS Escherichia coli.  
 OC Plasmid F.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 NCBI\_TaxID=562;  
 OX NCBI [1]  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=K12;  
 RX MEDLINE=90034191; PubMed=2680768;  
 RA Jalajakumari M.B., Manning P.A.;  
 RT "Nucleotide sequence of the trd region in the Escherichia coli F sex  
 RT factor."  
 RL gene 81:195-202(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94359430; PubMed=7915817;  
 RA Frost L.S., Ippen-Ihler K., Skurray R.A.;  
 RT Analysis of the sequence and gene products of the transfer region of  
 RT the F sex factor."  
 RL Microbiol. Rev. 58:162-210(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / CR63;  
 RA Shimizu H., Saitoh Y., Suda Y., Uehara K., Sempel G., Mizobuchi K.;  
 RT "Complete nucleotide sequence of the F plasmid: its implications for  
 RT organization and diversification of plasmid genomes."  
 RL submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 8-34 FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=88118903; PubMed=3323526;  
 RA Jalajakumari M.B., Guidolin A., Buhj H.J., Manning P.A.;  
 RT "Surface exclusion genes tras and trar of the F sex factor of  
 RT Escherichia coli K-12. Determination of the nucleotide sequence and  
 RT promoter and terminator activities."  
 RL J. Mol. Biol. 198:1-11(1987).  
 RN [5]  
 RP SEQUENCE OF 462-717 FROM N.A.  
 RX MEDLINE=90299847; PubMed=2163400;  
 RA Bradshaw H.D. Jr., Traxler B.A., Minkley E.G. Jr., Nester E.W.,  
 RA Gordon M.P.;  
 RT "Nucleotide sequence of the trar (hellicase I) gene from the sex  
 RT factor F."  
 RL J. Bacteriol. 172:4127-4131(1990).  
 RN [6]  
 RP SEQUENCE OF 1-102 AND 665-717 FROM N.A.  
 RX MEDLINE=90317835; PubMed=2164585;  
 RA Yoshioke Y., Fujita Y., Ohtsubo E.;  
 RT "Nucleotide sequence of the promoter-distal region of the tra operon  
 RT of plasmid R100, including tral (DNA helicase I) and trd genes."  
 RL J. Mol. Biol. 214:39-53(1990).  
 RN [7]  
 RP CHARACTERIZATION.  
 RX MEDLINE=92316963; PubMed=1618779;  
 RA Panicker M.M., Minkley E.G. Jr.;  
 RT "Purification and properties of the F sex factor Trad protein, an  
 RT inner membrane conjugal transfer protein."  
 RL J. Biol. Chem. 267:12761-12766(1992).  
 RN [8]  
 RP TOPOLOGY.  
 RX MEDLINE=94429860; PubMed=10498725;  
 RA Lee M.H., Kosuk N., Bailey J., Traxler B., Mancil C.;  
 RT "Analysis of F factor Trad membrane topology by use of gene fusions  
 RT and trypsin-sensitive insertions."  
 RL J. Bacteriol. 181:6108-6113(1999).  
 CC -1- FUNCTION: TRAD IS A CELL ENVELOPE PROTEIN. THE FUNCTION OF TRAD  
 CC IS UNKNOWN BUT IT IS ESSENTIAL FOR DNA TRANSFER. IT MAY FORM OR  
 CC MAY BE PART OF A NON-SPECIFIC PORE FOR DNA EXPORT, OR IT MAY  
 CC DIRECTLY ENERGIZES DNA TRANSPORT.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
 CC -1- SIMILARITY: TO THE TRAD OF PLASMID INCFIT R100.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

DR EMBL: M29254; AAA83928.1; -  
DR EMBL: U01159; AAC44181.1; -  
DR EMBL: AP001918; BAA97972.1; -  
DR EMBL: X06915; CAA30013.1; ALT\_SEQ.  
DR EMBL: M54796; AAA98083.1; -  
DR EMBL: X57431; CAA40678.1; -  
DR EMBL: X57428; CAA40674.1; -  
DR PIR: J50293; BVECAD.  
DR PIR: S01758; S01758.  
DR Ecocore: EG40117; trad.  
DR InterPro: IPR003688; TRAG.  
DR Pfam: PF02534; TRAG; 1.  
KW Plasmid; Inner membrane; Conjugation; Transmembrane; ATP-binding;  
KW Complete proteome.  
FT DOMAIN 1 27 CYTOPLASMIC (PROBABLE).  
FT TRANSMEM 28 47 PROBABLE.  
FT DOMAIN 48 104 PERIPLASMIC (PROBABLE).  
FT TRANSMEM 105 130 PROBABLE.  
FT DOMAIN 131 717 CYTOPLASMIC (PROBABLE).  
FT NP\_BIND 192 199 ATP (POTENTIAL).  
FT CONFLICT 19 47 IMMEDIATELY FOLLOWINGLY -> YPHVOPDR  
SQ SEQUENCE 717 AA; 81683 MW; F4B564EDD90EB914 CRC64;  
Query Match 86.1%; Score 31; DB 1; Length 717;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IFFWI 5  
DB 35 IFFWI 39  
RESULT 14  
TRD2\_ECOLI STANDARD; PRT; 738 AA.  
AC P22708;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
RT trad protein.  
RA Escherichia coli.  
OC Plasmid IncFII R100.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90317835; PubMed=2164585;  
RA Yoshida Y., Fujita Y., Ohtsuo E.;  
RT "Nucleotide sequence of the promoter-distal region of the tra operon  
of plasmid R100, including traI (DNA helicase I) and tra genes.";  
RL J. Mol. Biol. 214:39-53(1990).  
-1- FUNCTION: TRAD IS A CELL ENVELOPE PROTEIN. THE FUNCTION OF TRAD  
IS UNKNOWN BUT IT IS ESSENTIAL FOR DNA TRANSFER. IT MAY FORM OR  
MAY BE PART OF A NON-SPECIFIC PORE FOR DNA EXPORT, OR IT MAY  
DIRECTLY ENERGIZES DNA TRANSPORT.  
-1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
-1- SIMILARITY: TO THE TRAD OF PLASMID F.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: X55815; CAA39336.1; -  
DR PIR: S10659; S10659.  
DR InterPro: IPR003688; TRAG.  
DR Pfam: PF02534; TRAG; 1.  
KW Plasmid; Inner membrane; Conjugation; Transmembrane; ATP-binding;  
KW Repeat.  
FT TRANSMEM 28 44 POTENTIAL.  
FT TRANSMEM 113 129 POTENTIAL.  
FT TRANSMEM 396 413 POTENTIAL.  
FT NP\_BIND 192 199 ATP (POTENTIAL).  
FT DOMAIN 617 646 10 X 3 AA TANDEM REPEATS OF Q-Q-P.  
SQ SEQUENCE 738 AA; 83900 MW; 84CB1F48245E766F CRC64;  
Query Match 86.1%; Score 31; DB 1; Length 738;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IFFWI 5  
DB 35 IFFWI 39  
RESULT 15  
VILLI\_CHICK STANDARD; PRT; 826 AA.  
AC P02640;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-JUL-1989 (Rel. 11, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Villin.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88276884; PubMed=2839826;  
RA Bazzari W.L., Matsudaira P., Waller M., Smeal T., Jakes R., Ahmed Y.;  
RT "Villin sequence and peptide map identify six homologous domains.";  
RL Proc. Natl. Acad. Sci. U.S.A. 85:4986-4990(1988).  
RN [2]  
RP SEQUENCE OF 751-826.  
RX MEDLINE=81264203; PubMed=6790532;  
RA Glenney J.R., Jr., Geisler N., Kaulfus P., Weber K.;  
RT "Demonstration of at least two different actin-binding sites in  
RT villin, a calcium-regulated modulator of F-actin organization.";  
RL J. Biol. Chem. 256:8156-8161(1981).  
RN [3]  
RP CALCIUM-BINDING SITES.  
RX MEDLINE=83082892; PubMed=6848508;  
RA Hesterberg L.K., Weber K.;  
RT "Demonstration of three distinct calcium-binding sites in villin, a  
RT modulator of actin assembly.";  
RL J. Biol. Chem. 258:365-369(1983).  
RN [4]  
RP STRUCTURE BY NMR OF 1-127.  
RX MEDLINE=94191534; PubMed=8142900;  
RA Markus M.A., Nakayama T., Matsudaira P., Wagner G.;  
RT "Solution structure of villin 14T, a domain conserved among actin-  
RT seveling proteins.";  
RL Protein Sci. 3:70-81(1994).  
RN [5]  
RP STRUCTURE BY NMR OF 1-127.  
RX MEDLINE=97337440; PubMed=9194180;  
RA Markus M.A., Matsudaira P., Wagner G.;  
RT "Refined structure of villin 14T and a detailed comparison with other  
RT actin-severing domains.";  
RL Protein Sci. 6:1197-1209(1997).  
RN [6]

RP STRUCTURE BY NMR OF 792-826.  
 RX MEDLINE-97307248; PubMed-9164455;  
 RA McKnight C.J.; Matsudaira P.T.; Kim P.S.;  
 RT "NMR structure of the 35-residue villin headpiece subdomain.";  
 CC Nat. Struct. Biol. 4:180-184(1997).  
 CC -1- FUNCTION: CA(2+)-REGULATED ACTIN-BINDING PROTEIN. ITS CAPPING  
 CC FUNCTION IS LOCALIZED TO DOMAIN I; ITS SEVERING ACTION IS A  
 CC COOPERATION BETWEEN AN F-ACTIN BINDING SITE OF DOMAIN II AND THE  
 CC CAPPING ACTIVITY OF DOMAIN I.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- TISSUE SPECIFICITY: COMPONENT OF BRUSH BORDER MICROVILLI.  
 CC -1- SIMILARITY: BELONGS TO THE VILLIN/GELSOLIN FAMILY.  
 CC -1- SIMILARITY: CONTAINS 6 GELSOLIN-LIKE REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: J03781; AAA9133.1; -.  
 DR PIR: A03082; A03082.  
 DR PIR: A31822; A31822.  
 DR PDB: 2VIL; 01-APR-97.  
 DR PDB: 2VIL; 01-APR-97.  
 DR PDB: 1VII; 12-AUG-97.  
 DR InterPro: IPR001974; Gelsolin.  
 DR InterPro: IPR003128; VHP.  
 DR Pfam: PF00626; Gelsolin; 6.  
 DR Pfam: PF02209; VHP; 1.  
 DR PRINTS: PR00597; GELSOLIN.  
 DR SMART: SM00262; GEL; 6.  
 DR SMART: SM00153; VHP; 1.  
 DR Cytoskeleton; Calcium; Actin-binding; Capping protein; Repeat;  
 KW 3D-structure.  
 KM  
 FT DOMAIN 1 734 CORE.  
 FT REPEAT 735 826 HEADPIECE.  
 FT REPEAT 27 76 GELSOLIN-LIKE 1.  
 FT REPEAT 148 188 GELSOLIN-LIKE 2.  
 FT REPEAT 265 309 GELSOLIN-LIKE 3.  
 FT REPEAT 408 457 GELSOLIN-LIKE 4.  
 FT REPEAT 528 568 GELSOLIN-LIKE 5.  
 FT REPEAT 631 672 GELSOLIN-LIKE 6.  
 FT REPEAT 820 823 ABSOLUTEY REQUIRED FOR ACTIVITY.  
 FT SITE 129 137 CRUCIAL FOR BINDING AN ACTIN FILAMENT.  
 FT SITE 112 119 POLYPHOSPHOINOSITIDE BINDING (BY  
 FT SITE POLYPHOSPHOINOSITIDE BINDING (BY  
 FT SITE SIMILARITY).  
 FT SITE POLYPHOSPHOINOSITIDE BINDING (BY  
 FT SITE SIMILARITY).

FT STRAND 21 23  
 FT STRAND 30 32  
 FT TURN 34 36  
 FT STRAND 39 41  
 FT STRAND 44 44  
 FT STRAND 47 53  
 FT STRAND 58 64  
 FT STRAND 68 68  
 FT STRAND 72 88  
 FT TURN 89 90  
 FT STRAND 95 97  
 FT HELIX 104 110  
 FT TURN 112 113  
 FT STRAND 116 118  
 FT TURN 124 125  
 SQ SEQUENCE 826 AA; 92479 MW; 6A8898F7DF947389 CRC64;

Db 658 IFFWI 662  
 Search completed: January 3, 2003, 15:29:23  
 Job time: 6.18182 secs

Query Match 86.1%; Score 31; DB 1; Length 826;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 IFFWI 5

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:24:35 ; Search time 20.3182 Seconds  
(without alignments)  
60.846 Million cell updates/sec

Title: US-09-543-188a-23

Perfect score: 36  
Sequence: 1 IFFWK 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

al number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL\_21:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.virus:\*  
13: sp.vertibrate:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriap:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	35	97.2	707	5	095XL6	095XL6 caenorhabdi
2	34	94.4	158	2	093ST6	093ST6 chlorobium
3	34	94.4	409	2	0926D5	0926D5 haemophilus
4	33	91.7	62	12	091FW6	091FW6 chilo iride
5	33	91.7	101	16	09A628	09A628 caulobacter
6	33	91.7	580	2	0926G8	0926G8 vibrio para
7	32	88.9	73	8	036329	036329 chondrus cr
8	32	88.9	112	2	08VNI1	08VNI1 lactobacill
9	32	88.9	113	13	09PW03	09PW03 laboe rohit
10	32	88.9	120	16	09CMX2	09CMX2 pasteurella
11	32	88.9	174	2	0916G8	0916G8 lactobacill
12	32	88.9	174	2	08VNI6	08VNI6 lactobacill
13	32	88.9	423	16	098Q05	098Q05 mycoplasma
14	32	88.9	666	10	09FDV6	09FDV6 faqus sylvia
15	32	88.9	671	16	08UDJ6	08UDJ6 agrobacteri
16	32	88.9	819	12	091EJ5	091EJ5 human calic

17	32	88.9	849	12	093125	093125 human calic
18	32	88.9	988	16	083345	083345 treponema p
19	32	88.9	2280	12	08U280	08U280 human calic
20	31	86.1	32	9	09MCP6	09MCP6 bacterioph
21	31	86.1	74	8	09MTD6	09MTD6 toxoplasma
22	31	86.1	86	16	08ZMI0	08ZMI0 salmoneila
23	31	86.1	86	16	08Z496	08Z496 salmoneila
24	31	86.1	174	2	P97065	P97065 salmoneila
25	31	86.1	174	2	P97066	P97066 salmoneila
26	31	86.1	174	2	P97067	P97067 salmoneila
27	31	86.1	174	2	P97068	P97068 salmoneila
28	31	86.1	174	2	P97231	P97231 salmoneila
29	31	86.1	180	2	P71251	P71251 escherichia
30	31	86.1	180	2	P71252	P71252 escherichia
31	31	86.1	180	2	P71253	P71253 escherichia
32	31	86.1	180	2	P71254	P71254 escherichia
33	31	86.1	180	2	P71255	P71255 escherichia
34	31	86.1	180	2	P71256	P71256 escherichia
35	31	86.1	180	2	P71257	P71257 escherichia
36	31	86.1	180	2	P71258	P71258 escherichia
37	31	86.1	180	2	P71259	P71259 escherichia
38	31	86.1	180	2	P71260	P71260 escherichia
39	31	86.1	180	2	P71261	P71261 escherichia
40	31	86.1	180	2	P75018	P75018 escherichia
41	31	86.1	190	2	09K1A9	09K1A9 escherichia
42	31	86.1	190	2	09K1A8	09K1A8 escherichia
43	31	86.1	225	9	064371	064371 lactobacill
44	31	86.1	258	9	09A2P8	09A2P8 bacterioph
45	31	86.1	258	16	08ZS86	08ZS86 anabaena sp

#### ALIGNMENTS

##### RESULT 1

ID 095XL6 PRELIMINARY: PRT: 707 AA.  
AC 095XL6;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Hypothetical 79.9 KDa protein.  
GN Y3E7A.8.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Du H., Maupin R.;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Du H., Maupin R.;  
RX "The sequence of C. elegans cosmid Y3E7A.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Waterston R.;  
RX "Direct Submission.";  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC025727; AAK6029.1; -  
DR InterPro: IPR000884; TSP1.  
DR PROSITE: PSS0092; TSP1.1.  
KW Hypothetical protein.  
SQ SEQUENCE 707 AA; 79940 MW; 26B924988153DBA8 CRC64;

Query Match 97.2%; Score 35; DB 5; Length 707;  
Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6  
11111  
DB 470 IFFWIK 475

## RESULT 2

Q93ST6 PRELIMINARY; PRT; 158 AA.  
ID Q93ST6  
AC Q93ST6;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE BCHF.  
GN BCHF.  
OS Chlorobium tepidum.  
Bacteria: Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;  
Chlorobium.  
NCBI\_TaxID=1097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20433268; PubMed=10976061;  
RA Xiong J., Fischer W.M., Inoue K., Nakahara M., Bauer C.E.;  
RT "Molecular evidence for the early evolution of photosynthesis.";  
RL Science 289:1724-1730(2000).  
DR EMBL:AT005138; AAC12430.1; -  
SQ SEQUENCE 158 AA; 18396 MW; ABFEF8CF70B7411 CRC64;

Query Match 94.4%; Score 34; DB 2; Length 158;  
Best Local Similarity 83.3%; Pred. No. 72;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6  
11111  
DB 109 IFFWIK 114

## RESULT 3

Q926D5 PRELIMINARY; PRT; 409 AA.  
ID Q926D5;  
AC Q926D5;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)  
DE Ocapa.  
GN Haemophilus ducreyi1.  
OS Bacteria: Proteobacteria; gamma subdivision; Pasteurellaceae;  
Haemophilus.  
NCBI\_TaxID=730;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=35000;  
RC MEDLINE=99115572; PubMed=9916106;  
RA Bauer B.A., Lumblay S.R., Hansen E.J.;  
RT "Characterization of a Waaf (Rfaf) homolog expressed by Haemophilus  
ducreyi.";  
RL Infect. Immun. 67:899-907(1999).  
DR EMBL:AF087414; AAD16054.1; -  
SQ SEQUENCE 409 AA; 45059 MW; 2AC6861B1AC451 CRC64;

Query Match 94.4%; Score 34; DB 2; Length 409;  
Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6  
11111  
DB 161 IFFWIK 166

## RESULT 4

Q91FW6 PRELIMINARY; PRT; 62 AA.  
ID Q91FW6  
AC Q91FW6;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE 204L.  
OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).  
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.  
OX NCBI\_TaxID=10488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Delius H., Darai G., Fluegel R.M.;  
RT "DNA analysis of insect iridescent virus 6: evidence for circular  
permutation and terminal redundancy.";  
RL J. Virol. 49:609-614(1984).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86174607; PubMed=3959991;  
RA Lorbacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;  
RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis  
in mice.";  
RL Med. Microbiol. Immunol. 175:43-53(1986).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87321126; PubMed=2820141;  
RA Schmitzler P., Soltan J.B., Fischer M., Reiser H., Scholz J.,  
RA Delius H., Darai G.;  
RT "Molecular cloning and physical mapping of the genome of insect  
iridescent virus type 6: further evidence for circular permutation of  
the viral genome.";  
RL Virology 160:66-74(1987).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89073752; PubMed=3201750;  
RA Fischer M., Schmitzler P., Delius H., Darai G.;  
RT "Identification and characterization of the repetitive DNA element in  
the genome of insect iridescent virus type 6.";  
RL Virology 167:485-496(1988).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92196996; PubMed=1549908;  
RA Handermann M., Schmitzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,  
RA Darai G.;  
RT "Identification and mapping of origins of DNA replication within the  
DNA sequences of the genome of insect iridescent virus type 6.";  
RL Virus Genes 6:19-32(1992).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93118242; PubMed=1475907;  
RA Sonntag K.C., Darai G.;  
RT "Characterization of the third origin of DNA replication of the genome  
of insect iridescent virus type 6.";  
RL Virus Genes 6:333-342(1992).  
RN [7]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93260401; PubMed=8492091;  
RA Stohwasser R., Raab K., Schmitzler P., Janssen W., Darai G.;  
RT "Identification of the gene encoding the major capsid protein of  
insect iridescent virus type 6 by polymerase chain reaction.";  
RL J. Gen. Virol. 74:873-879(1993).  
RN [8]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94167241; PubMed=8121799;  
RA Schmitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,  
RA Delius H., Darai G.;  
RT "Identification of genes encoding zinc finger proteins, non-histone  
chromosomal HMG protein homologue, and a putative GTP phosphohydrolase  
in the genome of Chilo iridescent virus.";  
RL Nucleic Acids Res. 22:156-166(1994).  
RN [9]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94353641; PubMed=8073636;



```

RA Sonntag K.C., Schmitzler P., Koonin E.V., Darai G.;
RT "Chilo iridescent virus encodes a putative helicase belonging to a
RT distinct family within the 'DEAD/H' superfamily: implications for the
RT evolution of large DNA viruses.";
RL Virus Genes 8:151-158(1994).
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=95213160; PubMed=7698884;
RA Sonntag K.C., Schmitzler P., Janssen W., Darai G.;
RT "Identification of the primary structure and the coding capacity of
RT the genome of insect iridescent virus type 6 between the genome
RT coordinates 0.310 and 0.347 (7990 bp).";
RL Intervirology 37:287-297(1994).
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=94292906; PubMed=8021587;
RA Schmitzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,
RA Koonin E.V., Darai G.;
RT "Insect iridescent virus type 6 encodes a polypeptide related to the
RT largest subunit of eukaryotic RNA polymerase II.";
RL J. Gen. Virol. 75:1557-1567(1994).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=96141693; PubMed=9482589;
RA Bahr U., Tidona C.A., Darai G.;
RT "The DNA sequence of Chilo iridescent virus between the genome
RT coordinates 0.101 and 0.391: similarities in coding strategy between
RT insect and vertebrate iridoviruses.";
RL Virus Genes 15:235-245(1997).
RN [13]
RP SEQUENCE FROM N.A.
RX MEDLINE=99125223; PubMed=9926400;
RA Muller K., Tidona C.A., Bahr U., Darai G.;
RT "Identification of a thymidylate synthase gene within the genome of
RT Chilo iridescent virus.";
RL Virus Genes 17:243-258(1998).
RN [14]
RP SEQUENCE FROM N.A.
RX MEDLINE=99383793; PubMed=10456793;
RA Muller K., Tidona C.A., Darai G.;
RT "Identification of a gene cluster within the genome of Chilo
RT iridescent virus encoding enzymes involved in viral DNA replication
RT and processing.";
RL Virus Genes 18:243-264(1999).
RN [15]
RP SEQUENCE FROM N.A.
RX MEDLINE=21342589; PubMed=11448171;
RA Jakob N.J., Muller K., Bahr U., Darai G.;
RT "Analysis of the First Complete DNA Sequence of an Invertebrate
RT Iridovirus: Coding Strategy of the Genome of Chilo Iridescent Virus.";
RL Virology 286:182-196(2001).
RN [16]
RP SEQUENCE FROM N.A.
RA Jakob N.J., Muller K., Bahr U., Darai G.;
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF303741; AAK82066.1; -
SQ SEQUENCE 62 AA; 7876 MW; 29DF67A85F664B3E CRC64;

Query Match 91.7%; Score 33; DB 12; Length 62;
Best Local Similarity 83.3%; Pred. No. 46;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWIK 6
DB 56 IFFWIK 61

RESULT 5
O9A628 PRELIMINARY; PRT; 101 AA.
AC O9A628;
DT 01-JUN-2001 (TRENBLREL. 17, Created)
DR 01-JUN-2001 (TRENBLREL. 17, Last sequence update)

```

```

DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE Hypothetical protein CC2267.
GN CC2267.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19089 / CB15;
RA MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Givoni M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Kouri H., Shetty J., Berry K.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005897; AAK24238.1; -
DR TIGR: CC2267; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 101 AA; 11619 MW; 4B5B638942C94AF6 CRC64;

Query Match 91.7%; Score 33; DB 16; Length 101;
Best Local Similarity 66.7%; Pred. No. 72;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWIK 6
DB 6 VFFWIK 11

RESULT 6
O926G8 PRELIMINARY; PRT; 580 AA.
AC O926G8;
DT 01-MAY-1999 (TRENBLREL. 10, Created)
DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Polar flagellar M-ring protein FlIF.
GN FLIF.
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BB22;
RA MEDLINE=20115548; PubMed=10648530;
RA Boles B.R., McCarter L.L.;
RT "Insertional inactivation of genes encoding components of the sodium-
RT type flagellar motor and switch of vibrio parahaemolyticus.";
RL J. Bacteriol. 182:1035-1045(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=BB22;
RA Jaques S., Kim Y.K., McCarter L.L.;
RT "Components of the polar flagellar switch complex and assembly
RT apparatus.";
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF069392; AAD15920.1; -
DR InterPro: IPR000067; FlgM_ring_FLIF.
DR Pfam: PF01514; YscJ_FLIF; 1.
DR PRINTS: PRO1009; FLGMRINGFLIF.
DR TIGRFAMs: TIGR00206; flif; 1.
KW Flagella.
SQ SEQUENCE 580 AA; 63813 MW; 4CCE81483A047050 CRC64;

Query Match 91.7%; Score 33; DB 2; Length 580;
Best Local Similarity 66.7%; Pred. No. 3; 5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 IFFWK 6  
:||||:  
DB 71 IFFWK 76

## RESULT 7

Q36329 PRELIMINARY; PRT; 73 AA.  
AC Q36329;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Putative orf73 protein.  
GN PUTATIVE ORF73.  
OS Chondrus crispus (Carragheen).  
OC Mitochondrion.  
OC Eukaryota; Rhodophyta; Florideophyceae; Gigartinales; Gigartinales;  
OC Chondrus.  
NCBI\_TaxID=2769;  
[1]  
SEQUENCE FROM N.A.

RC TISSUE=APICES.  
RX MEDLINE=95341681; PubMed=7616569;  
RA Leblanc C., Boyen C., Richard O., Bonnard G., Grienenberger J.M.,  
RA Kioareg B.;  
RT "Complete sequence of the mitochondrial DNA of the rhodophyte Chondrus  
RT crispus (Gigartinales). Gene content and genome organization.";  
RL J. Mol. Biol. 250:484-495(1995).  
RN [2]  
RN SEQUENCE FROM N.A.

RC TISSUE=APICES;  
RX MEDLINE=95395875; PubMed=7666449;  
RA Leblanc C., Kioareg B., Loiseau-de Goer S., Boyen C.;  
RT "DNA sequence, structure and phylogenetic relationship of the  
RT mitochondrial small subunit rRNA from the red alga Chondrus crispus  
RT (Gigartinales, Rhodophytes).";  
RL J. Mol. Evol. 41:196-202(1995).  
DR EMBL: 247547; CAA87598.1; -.  
KM Mitochondrion.  
SQ SEQUENCE 73 AA; 8929 MW; 85AC46C006834292 CRC64;

Query Match 88.9%; Score 32; DB 8; Length 73;  
Best Local Similarity 83.3%; Pred. No. 81;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 IFFWK 6  
:||||:  
DB 38 IFFWK 43

## RESULT 8

Q8VN11 PRELIMINARY; PRT; 112 AA.  
AC Q8VN11;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Putative DNA binding protein.  
OS Lactobacillus delbrueckii (subsp. lactis).  
OC Plasmid pJBL2.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;  
OC Lactobacillaceae; Lactobacillus.  
OK NCBI\_TaxID=29397;  
RN [1]  
RN SEQUENCE FROM N.A.

RP STRAIN=JCL14;  
RA Bourniquet A.A., Casey M.G., Mollet B., Pridmore R.D.;  
RT "DNA sequence and functional analysis of Lactobacillus delbrueckii  
RT subsp. lactis plasmids PM42 and pJBL2.";  
RL Plasmid 0:0-0(0).  
DR EMBL: AJ421486; CAD13351.1; -.  
KM Plasmid.

SQ SEQUENCE 112 AA; 13105 MW; 2EB5A6C08F8460FF CRC64;  
Query Match 88.9%; Score 32; DB 2; Length 112;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWK 6  
:||||:  
DB 94 IFFWK 99

## RESULT 9

Q9PW03 PRELIMINARY; PRT; 113 AA.  
AC Q9PW03;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Rohu liver hypothetical protein (fragment).  
GN LVHPL.  
OS Labeo rohita (Indian major carp).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Labeo.  
OK NCBI\_TaxID=84645;  
RN [1]  
RN SEQUENCE FROM N.A.

RP TISSUE=LIVER;  
RC "Partial cDNA for a hypothetical protein from Labeo rohita liver.";  
RT Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ249232; CAB54039.1; -.  
KM Hypothetical protein.  
FT NON\_TER 1 1  
FT 113 113  
SQ SEQUENCE 113 AA; 13469 MW; 6E36D649DA8FFB08 CRC64;

Query Match 88.9%; Score 32; DB 13; Length 113;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 IFFWK 6  
:||||:  
DB 6 IFFWK 10

## RESULT 10

Q9CMX2 PRELIMINARY; PRT; 120 AA.  
AC Q9CMX2;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Hypothetical protein PM0679.  
GN PM0679.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Pasteurella.  
OK NCBI\_TaxID=747;  
RN [1]  
RN SEQUENCE FROM N.A.

RP STRAIN=PM70;  
RX MEDLINE=21145866; PubMed=11248100;  
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida pm70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
DR EMBL: AE006104; AAK02763.1; -.  
KM Hypothetical protein; Complete proteome.  
SQ SEQUENCE 120 AA; 14167 MW; 46308C480002FF64 CRC64;

Query Match 88.9%; Score 32; DB 16; Length 120;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FFWIK 6  
 Db 112 FFWIK 116

RESULT 11  
 Q9L6G8 PRELIMINARY; PRT; 174 AA.

AC 09L6G8  
 DT 01-OCT-2000 (TRENBLREL. 15, Created)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
 DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)  
 DE Putative transposase.  
 OS Lactobacillus delbrueckii (subsp. bulgaricus).  
 NC Lactobacillaceae; Lactobacillus.  
 NCBI\_TaxID=1585;

SEQUENCE FROM N.A.  
 Accarte M.A., Raya R.R.;  
 "Structural analysis of pLBB1, a cryptic plasmid from Lactobacillus delbrueckii subsp. bulgaricus."  
 RT Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF236060; AAF61726.1; -.  
 KW Plasmid.  
 SO SEQUENCE 174 AA; 19597 MW; 8DDB65BCDCDB556 CRC64;

Query Match 88.9%; Score 32; DB 2; Length 174;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWIK 6  
 Db 81 LFFWIK 86

RESULT 12  
 Q8VN06 PRELIMINARY; PRT; 174 AA.

AC 08VN06  
 DT 01-MAR-2002 (TRENBLREL. 20, Created)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last sequence update)  
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
 DE Hypothetical 19.6 kDa protein.  
 OS Lactobacillus delbrueckii (subsp. lactis).  
 NC Plasmid pM42.  
 Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales; Lactobacillaceae; Lactobacillus.  
 NCBI\_TaxID=29397;

SEQUENCE FROM N.A.  
 Bourinquel A.A., Casey M.G., Mollet B., Pridmore R.D.;  
 "DNA sequence and functional analysis of Lactobacillus delbrueckii subsp. lactis plasmids pM42 and pJBL2."  
 RT Plasmid 0:0-0(0).  
 RL EMBL; AJ421627; CAD15746.1; -.  
 KW Hypothetical protein; Plasmid.  
 SO SEQUENCE 174 AA; 19555 MW; 278D0FB317C001D9 CRC64;

Query Match 88.9%; Score 32; DB 2; Length 174;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWIK 6  
 Db 81 LFFWIK 86

RESULT 13  
 Q98QA5 PRELIMINARY; PRT; 423 AA.

AC Q98QA5;  
 DT 01-OCT-2001 (TRENBLREL. 18, Created)  
 DT 01-OCT-2001 (TRENBLREL. 18, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE Hypothetical protein MYPV\_4610.  
 GN MYPV\_4610.  
 OS Mycoplasma pulmonis.  
 NC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 NCBI\_TaxID=2107;

SEQUENCE FROM N.A.  
 STRAIN-UAB CTIP;  
 MEDLINE=21267165; PubMed=11353084;  
 RA Chambaud I., Hellig R., Ferris S., Barbe V., Samson D., Gailsson F., Moszer I., Dydvig K., Wroblewski H., Vlati A., Rocha E.P.C., Blanchard A.;  
 "The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis."  
 RT Nucleic Acids Res. 29:2145-2153(2001).  
 DR EMBL; AL445564; CAC13634.1; -.  
 DR Mypulist; MYPV\_4610; -.  
 DR InterPro; IPR004477; COMEC-N-term.  
 DR TIGRFAMs; TIGR00360; COMEC-N-term; 1.  
 KW Hypothetical protein; Complete proteome.  
 SO SEQUENCE 423 AA; 50819 MW; B4A027C3E9264F9E CRC64;

Query Match 88.9%; Score 32; DB 16; Length 423;  
 Best Local Similarity 66.7%; Pred. No. 4e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWIK 6  
 Db 390 LFFWIK 395

RESULT 14  
 Q9FDV6 PRELIMINARY; PRT; 666 AA.

AC 09FDV6  
 DT 01-MAR-2001 (TRENBLREL. 16, Created)  
 DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)  
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
 DE Protein kinase.

GN PK2.  
 OS Fagus sylvatica (Beechnut).  
 NC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eucots I; Fagales; Fagaceae; Fagus.  
 NCBI\_TaxID=28930;  
 SEQUENCE FROM N.A.  
 TISSUE=SEED;  
 RA Lorenzo O., Rodriguez D., Nicolas C., Nicolas G.;  
 "Characterization and expression of two protein kinase genes and EIN-3 like gene, which are regulated by ABA and GA3 in dormant Fagus sylvatica seeds."  
 RT (in) Black M., Bradford K.J., Vazquez-Ramos J. (eds.);  
 SEED BIOLOGY-ADVANCES AND APPLICATIONS 32, pp.329-340,  
 CAB International, Oxfordshire, UK (2000).

SEQUENCE FROM N.A.  
 TISSUE=SEED;  
 RA Lorenzo O.;  
 Thesis (2000). Fisiologia Vegetal, Facultad de Biologia, Universidad de Salamanca.  
 CC - SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AJ298993; CAC09581.1; -.  
 DR InterPro; IPR001064; Crystalin.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.

DR PRODOM: PD000001: Euk\_pkinase: 1.  
 DR SMART: SM00220; S\_TKC: 1.  
 DR SMART: SM00219; TYRC: 1.  
 DR PROSITE: PS00225; CRYSTALLIN\_BETAGAMMA: UNKNOWN\_1.  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM: 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST: 1.  
 KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.  
 SO SEQUENCE 666 AA; 74385 MW; DA586B1276259C97 CRC64;

Query Match 88.9%; Score 32; DB 10; Length 666;  
 Best Local Similarity 100.0%; Pred. No. 6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 FFWIK 6  
 |||||  
 DB 369 FFWIK 373

## RESULT 15

Q8UJD6 PRELIMINARY; PRT; 671 AA.

DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hypothetical protein Atu5541.  
 GN ATU5541 OR AGR\_PAT\_807.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Plasmid AT.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_TaxID=176299;  
 RN [1]

RP MEDLINE=21608550; PubMed=11743193;  
 RX MEDLINE=21608550; PubMed=11743193;

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
 Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
 Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
 Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 Kutyavin T., Levy R., Li W.-J., McClelland E., Palmieri A.,  
 Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
 Zhang S., Yoo H., Tao Y., Biddle P., Jung W., Krespan W., Perry M.,  
 Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 Nester E.W.;  
 RA "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 C58."  
 RT Science 294:2317-2323(2001).  
 RL [2]

SEQUENCE FROM N.A.

RX MEDLINE=21608551; PubMed=11743194;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Outollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Houniel K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,  
 RA Wollam C., Allinger M., Doughly D., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,  
 RA Clelo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 Agrobacterium tumefaciens C58."  
 RL Science 294:2323-2328(2001).  
 DR EMBL: AE008974; AAL6227.1; ALT\_INIT.  
 DR EMBL: AE007922; AAK90918.1; -  
 KW Hypothetical protein; Plasmid; Complete proteome.  
 SO SEQUENCE 671 AA; 72766 MW; 43EDB89528013482 CRC64;

Query Match 88.9%; Score 32; DB 16; Length 671;  
 Best Local Similarity 83.3%; Pred. No. 6.1e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWIK 6  
 |||||  
 DB 369 IFFWIE 374

Search completed: January 3, 2003, 15:32:03  
 Job time : 21.3182 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 3, 2003, 15:25:16 ; Search time 9 seconds  
(without alignments)  
19.615 Million cell updates/sec

Title: US-09-543-188A-23

Perfect score: 36

Sequence: 1 IFFWIK 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues  
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*  
  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	94.4	3287	2	US-08-477-451-7
2	30	83.3	218	2	US-08-336-031-4
3	30	83.3	218	5	PCT-US95-06725-4
4	30	83.3	411	2	US-08-336-031-2
5	30	83.3	411	2	US-08-902-853-7
6	30	83.3	411	5	PCT-US95-06725-2
7	30	83.3	423	1	US-08-844-064-7
8	30	83.3	423	1	US-09-009-433-7
9	30	83.3	470	4	US-08-068-392-2
10	30	83.3	470	4	US-08-396-988-2
11	30	83.3	470	4	US-09-391-104-26
12	30	83.3	553	1	US-08-475-894-2
13	30	83.3	553	1	US-08-484-710-2
14	30	83.3	553	2	US-08-484-709-2
15	30	83.3	553	2	US-08-474-697-2
16	30	83.3	827	4	US-08-669-286-11
17	30	83.3	827	4	US-09-469-253-11
18	30	83.3	827	4	US-09-642-146-11
19	30	83.3	876	1	US-08-785-429-2
20	30	83.3	876	3	US-08-996-621-2
21	30	83.3	883	2	US-08-953-492-2
22	30	83.3	888	2	US-09-134-001C-3032
23	29	80.6	82	4	US-09-257-179-114
24	29	80.6	142	4	US-08-678-194-8
25	29	80.6	142	4	US-08-890-011-8
26	29	80.6	142	4	US-09-262-724-8
27	29	80.6	270	4	US-09-083-305-19

28	29	80.6	505	3	US-08-747-221B-14	Sequence 14, Appl
29	29	80.6	505	4	US-09-005-051-14	Sequence 14, Appl
30	29	80.6	530	3	US-08-747-221B-53	Sequence 53, Appl
31	29	80.6	530	4	US-09-005-051-53	Sequence 53, Appl
32	29	80.6	550	3	US-08-747-221B-19	Sequence 19, Appl
33	29	80.6	550	3	US-08-747-221B-58	Sequence 58, Appl
34	29	80.6	550	4	US-09-005-051-19	Sequence 19, Appl
35	29	80.6	550	4	US-09-005-051-58	Sequence 58, Appl
36	29	80.6	776	4	US-09-165-396-3	Sequence 3, Appl
37	29	80.6	1298	1	US-08-222-616-33	Sequence 33, Appl
38	29	80.6	1298	1	US-08-340-011-2	Sequence 2, Appl
39	29	80.6	1298	3	US-08-901-710-2	Sequence 2, Appl
40	29	80.6	1298	4	US-08-446-648-33	Sequence 33, Appl
41	29	80.6	1298	5	PCT-US95-04228-33	Sequence 33, Appl
42	29	80.6	1362	2	US-08-874-678-33	Sequence 33, Appl
43	29	80.6	1362	3	US-08-643-839-33	Sequence 33, Appl
44	29	80.6	1362	4	US-09-348-886-33	Sequence 33, Appl
45	29	80.6	1363	1	US-08-340-011-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-08-477-451-7  
; Sequence 7, Application US/08477451  
; Patent No. 592865  
; GENERAL INFORMATION:  
; APPLICANT: Corvaci, Antonello  
; TITLE OF INVENTION: Helicobacter Pylori CagI Region  
; NUMBER OF SEQUENCES: 46  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/477,451  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McClung, Barbara G.  
; REGISTRATION NUMBER: 33,113  
; REFERENCE/DOCKET NUMBER: 0335.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 510-601-2708  
; TELEFAX: 510-655-3542  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3287 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: Single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-477-451-7

Query Match 94.4%; Score 34; DB 2; Length 3287;  
Best Local Similarity 83.3%; Pred. No. 5.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWIK 6  
Db 679 IFFWIK 684

RESULT 2

US-08-336-031-4  
; Sequence 4, Application US/08336031  
; Patent No. 5817782  
; GENERAL INFORMATION:  
; APPLICANT: Jazwinski, S. M.  
; TITLE OF INVENTION: LAG1: A GENE FOR INCREASING THE  
; TITLE OF INVENTION: LONGEVITY OF EUKARYOTES  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336,031  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/253,875  
; FILING DATE: 03-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 93032  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 218 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-336-031-4

Query Match 83.3%; Score 30; DB 2; Length 218;  
Best Local Similarity 80.0%; Pred. No. 2e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 IFFWI 5  
:||||  
118 VFFWI 122

RESULT 3  
PCT-US95-06725-4  
; Sequence 4, Application PC/TUS9506725  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: LAG1: A GENE FOR INCREASING THE  
; TITLE OF INVENTION: LONGEVITY OF EUKARYOTES  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/06725  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/253,875 & 08/336,031  
; FILING DATE: 03-JUN-1994 & 08-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 93032  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 218 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-06725-4

Query Match 83.3%; Score 30; DB 5; Length 218;  
Best Local Similarity 80.0%; Pred. No. 2e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 IFFWI 5  
:||||  
118 VFFWI 122

RESULT 4  
US-08-336-031-2  
; Sequence 2, Application US/08336031  
; Patent No. 5817782  
; GENERAL INFORMATION:  
; APPLICANT: Jazwinski, S. M.  
; TITLE OF INVENTION: LAG1: A GENE FOR INCREASING THE  
; TITLE OF INVENTION: LONGEVITY OF EUKARYOTES  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336,031  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/253,875  
; FILING DATE: 03-JUN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digiglio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 93032  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 411 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-336-031-2

Query Match 83.3%; Score 30; DB 2; Length 411;

Best Local Similarity 80.0%; Pred. No. 3.7e+02; Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5

DB 311 VFFWI 315

RESULT 5

US-08-902-853-7

; Sequence 7, Application US/08902853

; Patent No. 5945330

; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Corley, Neil C.

; APPLICANT: Shah, Purvi

; APPLICANT: Lal, Preeti

; TITLE OF INVENTION: HUMAN LONGEVITY-ASSURANCE PROTEIN HOMOLOGS

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/902,853

; FILING DATE: Herewith

; CLASSIFICATION: ?

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0345 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 411 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GenBank

; CLONE: 541568

; US-08-902-853-7

; Query Match 83.3%; Score 30; DB 2; Length 411;

; Best Local Similarity 80.0%; Pred. No. 3.7e+02;

; Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5

DB 311 VFFWI 315

RESULT 6

PCT-US95-06725-2

; Sequence 2, Application PC/TUS9506725

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: LAG1: A GENE FOR INCREASING THE

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Scully, Scott, Murphy &amp; Presser

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: United States

; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/06725

; FILING DATE:

; CLASSIFICATION:

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/253,875 &amp; 08/336,031

; FILING DATE: 03-JUN-1994 &amp; 08-NOV-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Digiglio, Frank S.

; REGISTRATION NUMBER: 31,346

; REFERENCE/DOCKET NUMBER: 93032

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 742-4343

; TELEFAX: (516) 742-4366

; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 411 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; PCT-US95-06725-2

; Query Match 83.3%; Score 30; DB 5; Length 411;

; Best Local Similarity 80.0%; Pred. No. 3.7e+02;

; Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5

DB 311 VFFWI 315

RESULT 7

US-08-844-064-7

; Sequence 7, Application US/08844064

; Patent No. 5747314

; GENERAL INFORMATION:

; APPLICANT: Lawlor, Elizabeth

; TITLE OF INVENTION: No. 5747314el Compounds

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SmithKline Beecham Corporation

; STREET: 709 Swedeland Road

; CITY: King of Prussia

; STATE: PA

; COUNTRY: USA

; ZIP: 19406-0939

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/844,064

; FILING DATE: 18-APR-1997

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 9607991.8  
FILING DATE: 18-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Glimm, Edward R.  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31458-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 423 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-844-064-7

Query Match 83.3%; Score 30; DB 1; Length 423;  
Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5  
DB 279 IFFWV 283

RESULT 8  
US-09-009-433-7  
Sequence 7, Application US/09009433  
Patent No. 6087142  
GENERAL INFORMATION:  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6087142el Compounds  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/009,433  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/844,064  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: 9607991.8  
FILING DATE: 18-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Glimm, Edward R.  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31458-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 423 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-009-433-7

Query Match 83.3%; Score 30; DB 3; Length 423;  
Best Local Similarity 80.0%; Pred. No. 3.8e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5  
DB 279 IFFWV 283

RESULT 9  
US-08-068-392-2  
Sequence 2, Application US/08068392  
Patent No. 6150152  
GENERAL INFORMATION:  
APPLICANT: Shapiro, Steven M.  
TITLE OF INVENTION: Human Macrophage Metalloproteinase  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM  
STREET: 800 N. Lindbergh Blvd.  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63167  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/068,392  
FILING DATE: 19930528  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyer, Scott J.  
REGISTRATION NUMBER: 25275  
REFERENCE/DOCKET NUMBER: 07-24(12406)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)694-3117  
TELEFAX: (314)694-5435  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 470 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-068-392-2

Query Match 83.3%; Score 30; DB 4; Length 470;  
Best Local Similarity 80.0%; Pred. No. 4.2e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFWIK 6  
DB 305 FFWLK 309

RESULT 10  
US-08-396-988-2  
Sequence 2, Application US/08396988  
Patent No. 6204043  
GENERAL INFORMATION:  
APPLICANT: Shapiro, Steven M.  
TITLE OF INVENTION: Human Macrophage Metalloproteinase  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SM  
STREET: 800 N. Lindbergh Blvd.  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63167



COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/396,988  
FILING DATE: 01-MAR-1995  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/068,392  
FILING DATE: 28-MAY-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyer, Scott J.  
REGISTRATION NUMBER: 25275  
REFERENCE/DOCKET NUMBER: .07-24(12406)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)694-3117  
TELEFAX: (314)694-5435  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 470 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-396-988-2

Query Match 83.3%; Score 30; DB 4; Length 470;  
Best Local Similarity 80.0%; Pred. NO. 4.2e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFWIK 6  
|||:1  
Db 305 FFWLK 309

RESULT 11  
US-09-391-104-26  
; Sequence 26, Application US/09391104  
; Patent No. 6399371  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Falduto, Michael T.  
; APPLICANT: Magnuson, Scott R.  
; APPLICANT: Morgan, Douglas W.  
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,  
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS  
; TITLE OF INVENTION: OF USING SAME  
; FILE REFERENCE: 6073.US.PI  
; CURRENT APPLICATION NUMBER: US/09/391,104  
; CURRENT FILING DATE: 1999-09-07  
; PRIORITY APPLICATION NUMBER: US 08/814,394  
; PRIORITY FILING DATE: 1997-03-11  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 26  
; LENGTH: 470  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-391-104-26

Query Match 83.3%; Score 30; DB 4; Length 470;  
Best Local Similarity 80.0%; Pred. NO. 4.2e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFWIK 6  
|||:1  
Db 305 FFWLK 309

RESULT 12  
US-08-475-894-2  
; Sequence 2, Application US/08475894

Patent No. 5641748  
; GENERAL INFORMATION:  
; APPLICANT: Yen-Ming Hsu  
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,894  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Louis Myers  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: BGP-191  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 553 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-475-894-2

Query Match 83.3%; Score 30; DB 1; Length 553;  
Best Local Similarity 80.0%; Pred. NO. 4.9e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5  
|||:1  
Db 68 IFFWV 72

RESULT 13  
US-08-484-710-2  
; Sequence 2, Application US/08484710  
; Patent No. 5656438  
; GENERAL INFORMATION:  
; APPLICANT: Yen-Ming Hsu  
; TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, Suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,710  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Louis Myers  
REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: BGP-190  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 553 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-484-710-2

Query Match 83.3%; Score 30; DB 1; Length 553;  
Best Local Similarity 80.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWI 5  
||||:  
68 IFFWV 72

RESULT 14  
US-08-484-709-2  
Sequence 2, Application US/08484709  
Patent No. 5837844  
GENERAL INFORMATION:  
APPLICANT: Yen-Ming Hsu  
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,709  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Louis Myers  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: BGP-192  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 553 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-484-709-2

Query Match 83.3%; Score 30; DB 2; Length 553;  
Best Local Similarity 80.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWI 5  
||||:  
Db 68 IFFWV 72

RESULT 15  
US-08-474-697-2

Sequence 2, Application US/08474697  
Patent No. 6171800  
GENERAL INFORMATION:  
APPLICANT: Yen-Ming Hsu  
TITLE OF INVENTION: THE CAIP-LIKE GENE FAMILY  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, Suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,697  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Louis Myers  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: BGP-193  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 553 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-474-697-2

Query Match 83.3%; Score 30; DB 4; Length 553;  
Best Local Similarity 80.0%; Pred. No. 4.9e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IFFWI 5  
||||:  
Db 68 IFFWV 72

Search completed: January 3, 2003, 15:34:31  
Job time : 10 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: January 3, 2003, 15:32:11 ; Search time 4.77273 Seconds  
(without alignments)  
23.825 Million cell updates/sec

Title: US-09-543-188A-23  
Perfect score: 36  
Sequence: 1 IFFWIK 6

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 117078 seqs, 18951520 residues

al number of hits satisfying chosen parameters: 117078

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	97.2	487	9 US-09-738-626-6481	Sequence 6481, App
2	31	86.1	53	10 US-09-864-761-40676	Sequence 40676, A
3	31	86.1	58	10 US-09-867-550-866	Sequence 866, App
4	31	86.1	79	10 US-09-205-658-325	Sequence 325, App
5	31	86.1	951	10 US-09-815-242-10465	Sequence 10465, A
6	31	86.1	951	10 US-09-815-242-13853	Sequence 13853, A
7	31	86.1	1256	10 US-09-935-291A-6	Sequence 6, App1
8	30	83.3	137	10 US-09-867-550-832	Sequence 832, App
9	30	83.3	312	10 US-09-393-634-51	Sequence 51, App1
10	30	83.3	445	9 US-09-987-598-177	Sequence 177, App
11	30	83.3	445	9 US-09-988-293A-177	Sequence 177, App
12	30	83.3	445	9 US-10-063-547-32	Sequence 32, App1
13	30	83.3	445	9 US-09-989-735-177	Sequence 177, App
14	30	83.3	445	9 US-09-990-444-177	Sequence 177, App
15	30	83.3	445	9 US-09-989-730-177	Sequence 177, App
16	30	83.3	445	9 US-09-990-436-177	Sequence 177, App
17	30	83.3	445	9 US-09-991-181-177	Sequence 177, App
18	30	83.3	445	9 US-09-993-687-177	Sequence 177, App
19	30	83.3	445	10 US-09-989-722-177	Sequence 177, App

20	30	83.3	445	10 US-09-989-723-177	Sequence 177, App
21	30	83.3	445	10 US-09-989-279-177	Sequence 177, App
22	30	83.3	445	10 US-09-989-727-177	Sequence 177, App
23	30	83.3	445	10 US-09-989-731-177	Sequence 177, App
24	30	83.3	445	10 US-09-989-732-177	Sequence 177, App
25	30	83.3	445	10 US-09-991-073-177	Sequence 177, App
26	30	83.3	445	10 US-09-990-442-177	Sequence 177, App
27	30	83.3	445	10 US-09-991-163-177	Sequence 177, App
28	30	83.3	445	10 US-09-993-604-177	Sequence 177, App
29	30	83.3	445	10 US-09-990-456-177	Sequence 177, App
30	30	83.3	445	10 US-09-989-721-177	Sequence 177, App
31	30	83.3	445	12 US-10-006-867-32	Sequence 32, App1
32	30	83.3	445	12 US-10-052-586-148	Sequence 148, App
33	30	83.3	470	9 US-09-920-455-219	Sequence 219, App
34	30	83.3	470	10 US-09-801-196-23	Sequence 23, App1
35	30	83.3	473	10 US-09-925-301-1217	Sequence 1217, App
36	30	83.3	816	10 US-09-815-242-12636	Sequence 12636, A
37	30	83.3	876	10 US-09-815-242-13003	Sequence 13003, A
38	30	83.3	877	10 US-09-815-242-4876	Sequence 4876, App
39	30	83.3	879	10 US-09-815-242-5312	Sequence 5312, App
40	30	83.3	880	10 US-09-815-242-10675	Sequence 10675, A
41	30	83.3	883	10 US-09-815-242-13382	Sequence 13382, A
42	30	83.3	883	10 US-09-815-242-13684	Sequence 13684, A
43	30	83.3	950	10 US-09-815-242-11961	Sequence 11961, A
44	30	83.3	954	10 US-09-815-242-11238	Sequence 11238, A
45	29	80.6	34	10 US-09-864-761-49024	Sequence 49024, A

## ALIGNMENTS

RESULT 1  
US-09-738-626-6481  
; Sequence 6481, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738, 626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 6481  
; LENGTH: 487  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-6481

Query Match 97.2% ; Score 35; DB 9; Length 487;  
Best Local Similarity 83.3% ; Pred. No. 33;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWIK 6  
DB 412 VFWIK 417

```
RESULT 2
US-09-864-761-40676
; Sequence 40676, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-07-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40676
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC015473.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
; OTHER INFORMATION: SWISSPROT HIT: O02791, EVALUATE 1.00e-25
; OTHER INFORMATION: EST_HUMAN HIT: AW398979.1, EVALUATE 2.00e-24
US-09-864-761-40676

Query Match      86.1%; Score 31; DB 10; Length 53;
Best Local Similarity 100.0%; Pred. No. 22;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
```

```
Oy      1 IFFWI 5
      |||||
Db      38 IFFWI 42

RESULT 3
US-09-867-550-866
; Sequence 866, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fued
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells a
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 866
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-866

Oy      2 FFWIK 6
      |||||
Db      42 FFWIK 46

RESULT 4
US-09-205-658-325
; Sequence 325, Application US/09205658
; Patent No. US20010029617A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Ogg, Scott
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; FILE REFERENCE: 00786/351004
; CURRENT APPLICATION NUMBER: US/09/205,658
; CURRENT FILING DATE: 1998-12-03
; EARLIER APPLICATION NUMBER: 08/857,076
; EARLIER FILING DATE: 1997-05-15
; EARLIER APPLICATION NUMBER: 08/888,534
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: US98/10080
; EARLIER FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 328
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 325
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo sapiens or Caenorhabditis elegans
US-09-205-658-325

Query Match      86.1%; Score 31; DB 10; Length 79;
Best Local Similarity 83.3%; Pred. No. 31;
Matches      5; Conservative      0; Mismatches      1; Indels      0; Gaps      0;
```

```
RESULT 5
US-09-815-242-10465
; Sequence 10465, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10465
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10465

Query Match
Best Local Similarity 86.1%; Score 31; DB 10; Length 951;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5
    |||||
    512 IFFWI 516

RESULT 6
US-09-815-242-13853
; Sequence 13853, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```

```
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13853
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Salmoneella typhi
US-09-815-242-13853

Query Match
Best Local Similarity 86.1%; Score 31; DB 10; Length 951;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5
    |||||
    512 IFFWI 516

RESULT 7
US-09-935-291A-6
; Sequence 6, Application US/09935291A
; Patent No. US20020107373A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: 49937, 49931, AND 49933, NOVEL HUMAN TRANSPORTER
; FILE REFERENCE: MNI-184
; CURRENT APPLICATION NUMBER: US/09/935,291A
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/226,504
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: 60/250,932
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1256
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-291A-6

Query Match
Best Local Similarity 86.1%; Score 31; DB 10; Length 1256;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 FFWK 6
    ||||:|
    1025 FFWK 1029

RESULT 8
US-09-867-550-832
; Sequence 832, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehrahan, Ruad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells a
; FILE REFERENCE: 21402-013 (Cura-313)
```

```

: CURRENT APPLICATION NUMBER: US/09/867,550
: CURRENT FILING DATE: 2001-09-20
: PRIOR APPLICATION NUMBER: USSN 60/208,427
: PRIOR FILING DATE: 2000-05-30
: NUMBER OF SEQ ID NOS: 2125
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 832
: LENGTH: 137
: TYPE: PR
: ORGANISM: Homo sapiens
US-09-867-550-832
```

```

Query Match      83.3%; Score 30; DB 10; Length 137;
Best Local Similarity 80.0%; Pred. No. 75;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      1 1FFWI 5
        :1111
Db      130 VFFWI 134
```

```

RESULT 9
US-09-393-634-51
: Sequence 51, Application US/09393634
: Patent No. US20020051997A1
: GENERAL INFORMATION:
: APPLICANT: Zuker, Charles S.
: APPLICANT: Adler, Jon Elliot
: APPLICANT: Ryba, Nick
: APPLICANT: Mueller, Ken
: APPLICANT: Hoon, Mark
: APPLICANT: The Regents of the University of California
: APPLICANT: as represented by the United States of America
: APPLICANT: Department of Health and Human Services
: TITLE OF INVENTION: SF, a No. US20020051997A1el Family of Taste Receptors
: FILE REFERENCE: 02307E-098000US
: CURRENT APPLICATION NUMBER: US/09/393,634
: CURRENT FILING DATE: 1999-09-10
: NUMBER OF SEQ ID NOS: 92
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 51
: LENGTH: 312
: TYPE: PR
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: human GR09
US-09-393-634-51
```

```

Query Match      83.3%; Score 30; DB 10; Length 312;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      2 2FFWK 6
        :1111
Db      119 FFWLK 123
```

```

RESULT 10
US-09-992-598-177
: Sequence 177, Application US/09992598
: Patent No. US20020160384A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi J.
: APPLICANT: Baker, Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan L.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Fong, Sherman
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerltusen, Mary E.
: APPLICANT: Goddard, Audrey
```

```

: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2730PIC20
: CURRENT APPLICATION NUMBER: US/09/992,598
: CURRENT FILING DATE: 2001-11-14
: PRIOR APPLICATION NUMBER: 60/049787
: PRIOR FILING DATE: 1997-06-16
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/065186
: PRIOR FILING DATE: 1997-11-12
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066770
: PRIOR FILING DATE: 1997-11-24
: PRIOR APPLICATION NUMBER: 60/075945
: PRIOR FILING DATE: 1998-02-25
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/083322
: PRIOR FILING DATE: 1998-04-28
: PRIOR APPLICATION NUMBER: 60/084600
: PRIOR FILING DATE: 1998-05-07
: PRIOR APPLICATION NUMBER: 60/087106
: PRIOR FILING DATE: 1998-05-28
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087609
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087759
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087827
: PRIOR FILING DATE: 1998-06-03
: PRIOR APPLICATION NUMBER: 60/088021
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088025
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088026
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088028
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088029
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088030
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088033
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088326
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088167
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088202
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088212
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088217
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088655
: PRIOR FILING DATE: 1998-06-09
```

PRIOR APPLICATION NUMBER: 60/088734  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088738  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088742  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088824  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088826  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088861  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088876  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089105  
 PRIOR FILING DATE: 1998-06-12  
 PRIOR APPLICATION NUMBER: 60/089440  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089512  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089514  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089538  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089598  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089600  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089653  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089801  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089908  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089948  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089952  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090246  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090252  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090254  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090355  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090431  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090435  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090444  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090472  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090535

PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090540  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090542  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090557  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090676  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090678  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090690  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090694  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090695  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090696  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090862  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/090863  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/091360  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091478  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091544  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091519  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091626  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091633  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091978  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/091982  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/092182  
 PRIOR FILING DATE: 1998-07-09

Query Match 83.3%; Score 30; DB 9; Length 445;  
 Best Local Similarity 80.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IFFWI 5  
 Db 133 IFFWI 137

RESULT 11  
 US-09-989-293A-177  
 Sequence 177, Application US/09989293A  
 Patent No. US20020177164A1  
 GENERAL INFORMATION:  
 APPLICANT: Ashkenazi, Avi J.  
 APPLICANT: Baker, Kevin P.  
 APPLICANT: Botstein, David  
 APPLICANT: Desnoyers, Luc  
 APPLICANT: Baton, Dan L.  
 APPLICANT: Ferrara, Napoleone  
 APPLICANT: Fong, Sherman  
 APPLICANT: Geider, Hanspeter  
 APPLICANT: Gerlitsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Grimaldi, J. Christopher  
 APPLICANT: Gueney, Austin L.  
 APPLICANT: Kijavini, Ivar J.  
 APPLICANT: Napier, Mary A.  
 APPLICANT: Pan, James  
 APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730PIC66  
CURRENT APPLICATION NUMBER: US/09/989,293A  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810

PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24



PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Query Match 83.3%; Score 30; DB 9; Length 145;  
Best Local Similarity 80.0%; Pred. No. 2,1e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWT 5  
1111:  
DB 133 IFFWV 137

RESULT 12  
US-10-063-547-32  
Sequence 32, Application US/10063547  
Publication No. US20020182638A1

GENERAL INFORMATION:

APPLICANT: Eaton, Dan L.  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Geriltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3230R1C1  
CURRENT APPLICATION NUMBER: US/10/063,547  
PRIOR FILING DATE: 2002-05-02  
Prior Application removed - See file wrapper or Palm  
NUMBER OF SEQ ID NOS: 170  
SEQ ID NO 32  
LENGTH: 445  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-063-547-32

Query Match 83.3%; Score 30; DB 9; Length 145;  
Best Local Similarity 80.0%; Pred. No. 2,1e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IFFWT 5  
1111:  
DB 133 IFFWV 137

RESULT 13  
US-09-989-735-177  
Sequence 177, Application US/09989735  
Publication No. US20020193299A1  
GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Bolstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Geriltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C61  
CURRENT APPLICATION NUMBER: US/09/989,735  
PRIOR FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04

;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088028  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088029  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088030  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088033  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088326  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088167  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088202  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088212  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088217  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088655  
;; PRIOR FILING DATE: 1998-06-09  
;; PRIOR APPLICATION NUMBER: 60/088734  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088738  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088742  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088824  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088826  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089440  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089600  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246

;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 83.3%; Score 30; DB 9; Length 445;  
Best Local Similarity 80.0%; Pred. No. 2.1e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IFFWI 5  
DB 133 IFFWV 137

RESULT 14  
US-09-990-444-177  
Sequence 177, Application US/09990444  
Publication No. US20020193300A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730PIC19  
CURRENT APPLICATION NUMBER: US/09/990,444  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030

PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23

```

PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match      83.3%  Score 30; DB 9; Length 445;
Best Local Similarity 80.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1  IFFWI 5
      133  IFFWI 137

RESULT 15
US-09-989-730-177
; Sequence 177, Application US/09989730
; Publication No. US20020197674A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C69
CURRENT APPLICATION NUMBER: US/09/989,730
PRIOR FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
```

PRIOR APPLICATION NUMBER: 60/088202  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088212  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088217  
 PRIOR FILING DATE: 1998-06-05  
 PRIOR APPLICATION NUMBER: 60/088655  
 PRIOR FILING DATE: 1998-06-09  
 PRIOR APPLICATION NUMBER: 60/088734  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088738  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088742  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088810  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088824  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088826  
 PRIOR FILING DATE: 1998-06-10  
 PRIOR APPLICATION NUMBER: 60/088858  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088861  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/088876  
 PRIOR FILING DATE: 1998-06-11  
 PRIOR APPLICATION NUMBER: 60/089105  
 PRIOR FILING DATE: 1998-06-12  
 PRIOR APPLICATION NUMBER: 60/089440  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089512  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089514  
 PRIOR FILING DATE: 1998-06-16  
 PRIOR APPLICATION NUMBER: 60/089532  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089538  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089598  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089599  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089600  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089653  
 PRIOR FILING DATE: 1998-06-17  
 PRIOR APPLICATION NUMBER: 60/089801  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089907  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089908  
 PRIOR FILING DATE: 1998-06-18  
 PRIOR APPLICATION NUMBER: 60/089947  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089948  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/089952  
 PRIOR FILING DATE: 1998-06-19  
 PRIOR APPLICATION NUMBER: 60/090246  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090252  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090254  
 PRIOR FILING DATE: 1998-06-22  
 PRIOR APPLICATION NUMBER: 60/090349  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090355  
 PRIOR FILING DATE: 1998-06-23  
 PRIOR APPLICATION NUMBER: 60/090429  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090431  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090435

PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090444  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090445  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090472  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090535  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090540  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090542  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090557  
 PRIOR FILING DATE: 1998-06-24  
 PRIOR APPLICATION NUMBER: 60/090676  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090678  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090690  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090694  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090695  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090696  
 PRIOR FILING DATE: 1998-06-25  
 PRIOR APPLICATION NUMBER: 60/090862  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/090863  
 PRIOR FILING DATE: 1998-06-26  
 PRIOR APPLICATION NUMBER: 60/091360  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091478  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091544  
 PRIOR FILING DATE: 1998-07-01  
 PRIOR APPLICATION NUMBER: 60/091519  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091626  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091633  
 PRIOR FILING DATE: 1998-07-02  
 PRIOR APPLICATION NUMBER: 60/091978  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/091982  
 PRIOR FILING DATE: 1998-07-07  
 PRIOR APPLICATION NUMBER: 60/092182  
 PRIOR FILING DATE: 1998-07-09

Query Match 83.3%; Score 30; DB 9; Length 445;  
 Best Local Similarity 80.0%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 IFFWI 5  
 DB 133 IFFWI 137

Search completed: January 3, 2003, 15:52:07  
 Job time : 5.77273 secs

**THIS PAGE BLANK (USPTO)**